Journal of Mechanical Engineering and Manufacturing
https://www.sciltp.com/journals/jmem

Review

Check for updates

Machine Learning for Process Optimization and Defect Detection in Metal Additive Manufacturing: A Critical Review of Algorithms, *In-Situ* Monitoring Strategies, and Quality Assurance Frameworks

Ignatius Ekengwu *, Dara Jude, Anthonia Ilechukwu, Sunday Chimezie Anyaora and Augustine Uzochukwu Mmadumere

Department of Mechanical Engineering, Faculty of Engineering, Nnamdi Azikiwe University, Awka 420007, Anambra State, Nigeria
* Correspondence: ie.ekengwu@unizik.edu.ng

How To Cite: Ekengwu, I.; Jude, D.; Ilechukwu, A.; et al. Machine Learning for Process Optimization and Defect Detection in Metal Additive Manufacturing: A Critical Review of Algorithms, *In-Situ* Monitoring Strategies, and Quality Assurance Frameworks. *Journal of Mechanical Engineering and Manufacturing* **2026**. https://doi.org/10.53941/jmem.2026.100022

Received: 8 March 2026
Revised: 27 March 2026
Accepted: 30 March 2026
Published: 2 June 2026

Abstract: Among the many manufacturing technologies to emerge in the last three decades, metal additive manufacturing (AM)—spanning laser powder bed fusion (LPBF), directed energy deposition (DED), and electron beam melting (EBM)—stands out for the speed and breadth of its industrial adoption. Once confined to producing plastic concept models, it now fabricates flight-critical aerospace structures, load-bearing skeletal implants, and thermally demanding heat management components. Yet persistent process-induced defects—among them gas porosity, lack-of-fusion voids, hot cracking, and inter-layer delamination—continue to jeopardize part integrity and create formidable barriers to certification in regulated sectors. Machine learning (ML) and deep learning (DL) have attracted growing attention as remedies, offering the capacity to extract decision-relevant patterns from the rich multi-modal data streams that flow continuously during metal AM builds. This review undertakes a critical examination of the current state of ML-driven process optimization and defect detection for metal AM. Convolutional neural networks (CNNs) are evaluated for image-based anomaly identification; long short-term memory (LSTM) networks and transformer architectures are assessed for temporal process monitoring; reinforcement learning is examined for closed-loop parameter governance; and generative adversarial networks are considered as instruments for training data augmentation. *In-situ* monitoring hardware—encompassing melt-pool cameras, pyrometers, acoustic emission sensors, and photodiode arrays—is surveyed alongside the sensor fusion strategies that sharpen defect localization performance. Closed-loop quality assurance architectures, digital twin integration, and ML-guided parameter optimization receive detailed critical evaluation. Drawing on a systematic synthesis of more than 130 peer-reviewed studies spanning 2016 to 2025, we identify structural gaps: a shortage of large annotated benchmark datasets, insufficient cross-domain generalization, and the computational demands of real-time industrial deployment. A forward-looking research roadmap—centred on physics-informed neural networks, federated learning architectures, and edge-optimized inference—is advanced as the route toward certifiable, industrially viable ML quality assurance systems for metal AM.

Keywords: additive manufacturing; machine learning; defect detection; *in-situ* monitoring; process optimization; quality assurance

Copyright: © 2026 by the authors. This is an open access article under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).
Publisher's Note: Scilight stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

Few manufacturing technologies have transitioned from laboratory curiosity to industrial workhorse as quickly as metal additive manufacturing. In barely two decades, what began as a technique for producing plastic concept models has become a trusted production method for components operating at the edges of physical performance—turbine blades at altitude, orthopaedic implants under skeletal load, and combustor liners at temperatures that would destroy most materials. The geometry it makes possible is genuinely revolutionary: internal lattices, conformal cooling passages, and topology-optimised load paths that are not merely difficult but physically impossible to achieve through any conventional manufacturing route. Market projections indicate that global metal AM revenues will exceed USD 18 billion by 2030, driven by aerospace, defence, biomedical, and energy sectors that have collectively concluded that the technology's capabilities are worth its costs [1].

This commercial ascent has, however, never been free of a fundamental quality problem that anyone working seriously in the field recognises immediately. Unlike a machined part, where surface quality and dimensional accuracy are directly observable, a metal AM component can look perfect while harbouring internal damage that will eventually cause it to fail. A minute drift in laser power, a slightly contaminated powder particle, or a subtle change in shielding gas flow can introduce sub-surface pores, cracks, or unbonded inter-layer interfaces invisible to the naked eye yet capable of initiating catastrophic fatigue fracture under service conditions. The standard response—post-build inspection using X-ray computed tomography, ultrasonic examination, or destructive metallographic sectioning—is expensive, slow, and inherently reactive. By the time a defect is found, the part is finished and the only economically rational option is scrap. In high-value, low-volume applications—precisely the niche where metal AM has made its deepest inroads—this reactive quality logic is both costly and, in safety-critical contexts, unacceptable [2,3].

Machine learning reframes this problem entirely. Instead of examining the completed part, an ML-enabled system watches the build as it unfolds—learning to recognise the subtle precursor signatures of defect formation in thermal gradients, acoustic emissions, and melt-pool shape changes that develop milliseconds before a void or crack fully forms. Patterns that are detected earlier can be identified more reliably and at a fraction of the cost of post-process inspection. More compellingly, when a monitoring system is integrated with the machine controller through a real-time feedback loop, early detection enables active intervention: laser power is modulated, scan speed is adjusted, or hatch spacing is altered on the fly to suppress a developing anomaly before it consolidates into a defect. The concept of a manufacturing system capable of recognising and correcting its own errors has been discussed in the AM literature for well over a decade [4]. What is genuinely new is that it is now beginning to work.

Several enabling developments have converged to make this possible. Deep learning software frameworks have become accessible enough that researchers without dedicated ML expertise can implement and benchmark state-of-the-art architectures. The hardware required to support real-time inference—GPU accelerators that once required a server room—now fits into form factors compatible with machine-floor integration. Equally consequential, the research community has begun constructing the annotated training datasets that credible ML models require: built through expensive but necessary X-ray CT campaigns, systematic acoustic characterisation experiments, and purpose-designed monitoring test rigs that generate the volume and variety of labelled examples that shallow-dataset studies could never provide.

The critical evaluation that follows neither dismisses the genuine progress that has been made nor accepts published accuracy claims at face value. Our central question throughout is not whether an ML model achieved impressive results on its training domain, but whether those results would survive contact with real production conditions—a different machine, a different alloy batch, a different technician, or a build geometry not represented in the training data. We also examine what these algorithms are actually learning—whether internal representations correspond to physical reality or to artefacts of laboratory-controlled data collection—and where the technical, computational, and regulatory barriers lie between current capability and the certified quality assurance systems that industry genuinely needs. The analysis is restricted to metal AM processes—LPBF, DED, and EBM—which together present the most urgent industrial challenges and have attracted the largest share of ML research investment. The evidential basis is a systematic search across Web of Science, Scopus, and Google Scholar, yielding a synthesised corpus of 133 peer-reviewed primary sources from 2016 through 2025. Parallel comprehensive reviews of in-process monitoring confirm the maturity of the sensor hardware ecosystem while independently arriving at the same conclusion regarding the persistence of generalisation, data, and certification barriers [5].

2. Methodology

2.1. Literature Search Strategy and PRISMA Framework

The systematic review protocol adopted here follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines [6,7]. The PRISMA framework was selected because it provides a transparent, reproducible, and internationally recognised structure for identifying, screening, and adjudicating the inclusion of primary literature—qualities that matter particularly in interdisciplinary engineering reviews where research quality and indexing vary considerably across source types. Beyond methodological rigour, adherence to PRISMA has become an indexing requirement for an increasing number of journals, including those seeking entry into Scopus and the Web of Science Core Collection, and constitutes accepted best practice for review contributions to engineering and manufacturing disciplines.

2.2. Information Sources and Indexing Requirements

The search drew on three principal scholarly databases: Scopus (Elsevier), Web of Science Core Collection (Clarivate), and Google Scholar. Scopus and Web of Science were designated primary sources because they represent the most rigorously curated and comprehensively indexed multidisciplinary repositories currently available, and inclusion in these databases serves as the principal quality benchmark applied by the Journal of Mechanical Engineering and Manufacturing (JMEM) and its cognate publications when evaluating cited work. Accordingly, only records indexed in Scopus or Web of Science were eligible as primary evidence in the synthesis; Google Scholar was employed strictly as a supplementary discovery tool for grey literature, preprints, and conference proceedings whose formal indexing status required subsequent verification. Authoritative technical and regulatory documents from ISO, ASTM International, NIST, the U.S. Food and Drug Administration, and the Federal Aviation Administration were included where they constituted primary references that peer-reviewed journal articles could not adequately replace.

2.3. Search Terms and Time Period

The literature search was bounded by publication dates between January 2016 and December 2025. This window was chosen deliberately to capture the period during which deep learning methods became practically deployable for AM quality assurance—broadly coinciding with the maturation of GPU-accelerated training platforms and the availability of AM process datasets of sufficient scale and diversity to support genuine model learning. Work published before 2016 was consulted selectively for foundational context—process physics, classical inspection methodology—but was not subject to full PRISMA eligibility assessment. The primary Boolean search string, applied consistently across all three databases, combined AM process terms (“additive manufacturing”, “laser powder bed fusion”, “LPBF”, “selective laser melting”, “SLM”, “directed energy deposition”, “DED”, “electron beam melting”, “EBM”) with ML method terms (“machine learning”, “deep learning”, “neural network”, “convolutional neural network”, “CNN”, “reinforcement learning”, “transfer learning”, “generative adversarial network”, “GAN”, “LSTM”, “transformer”) and quality outcome terms (“defect detection”, “process monitoring”, “quality assurance”, “*in-situ* monitoring”, “process optimization”, “anomaly detection”). Supplementary targeted searches covered sensor fusion, digital twins, physics-informed neural networks, Bayesian optimisation, and federated learning applications in manufacturing contexts.

2.4. Inclusion and Exclusion Criteria

Eligibility assessment proceeded in two stages. During title and abstract screening, records were evaluated for relevance to the intersection of metal AM processes and machine learning or data-driven methods. Studies were excluded at this stage if they: (i) addressed polymer, ceramic, or composite AM without reference to metal processes; (ii) used only classical statistical or rule-based methods without any ML component; (iii) were computational materials science investigations lacking AM process context; or (iv) were duplicates, editorials, retractions, or non-peer-reviewed secondary sources. In the full-text stage, retained records were evaluated against four inclusion criteria: (a) the study must report original experimental or computational results on metal AM; (b) at least one ML algorithm must be applied for process monitoring, defect detection, quality assurance, or parameter optimisation; (c) quantitative performance metrics must be reported (accuracy, precision, recall, F1-score, RMSE, or equivalent); and (d) the full text must be accessible for detailed extraction. Studies presenting only qualitative assessments were excluded from the primary synthesis but may be referenced for contextual framing.

2.5. PRISMA Filtering and Data Extraction

The PRISMA cascade proceeded as follows. Initial database searches returned 4218 records (Scopus: 1847; Web of Science: 1623; Google Scholar supplementary: 748). Automated deduplication in Mendeley Reference Manager reduced this to 2971 unique records. Title and abstract screening identified 623 as potentially eligible. Full-text retrieval succeeded for 591; the remaining 32, despite inter-library loan requests, were excluded for inaccessibility. Full-text review resulted in the exclusion of a further 458 records: not specific to metal AM (n = 187); ML methods not meeting the defined criteria (n = 134); quantitative performance metrics absent (n = 76); review articles without primary data (n = 41); and results duplicating already-included primary studies (n = 20). The resulting 133 primary studies constitute the evidential base of this review [8,9], summarised in accordance with the PRISMA 2020 flow diagram framework. For each included study, the following data were extracted into a structured spreadsheet: AM process type; material(s); ML algorithm(s) employed; *in-situ* sensor modalities; performance metrics; dataset size; whether cross-machine or cross-material generalisation was tested; and whether real-time or embedded deployment was evaluated. This structured extraction underpins the quantitative comparisons in Section 3 and the critical assessment in Section 7.

3. Metal AM Processes and the Physics of Defect Formation

3.1. A Brief Process Taxonomy

To appreciate why machine learning is needed in metal AM, one must first understand the physical complexity of the processes it aims to oversee. Laser Powder Bed Fusion (LPBF) has attracted by far the greatest concentration of ML research, and the reasons are not coincidental: its layer-by-layer build architecture, standardised chamber geometry, and relatively accessible optical monitoring interfaces make it well-suited to the systematic, repeatable data collection that ML model training demands [10,11]. In LPBF, a high-power fibre laser—typically rated between 200 and 1000 W at a wavelength of 1064 nm—raster-scans across a freshly spread powder layer that is 20 to 100 µm thick and composed of particles ranging from 15 to 50 µm in diameter. The laser melts and fuses the powder along each scan track; the build platform then descends by one layer increment, a fresh powder layer is deposited, and the cycle repeats hundreds or thousands of times until the part is complete. Throughout this process, the build chamber is continuously purged with inert shielding gas—argon or nitrogen—to prevent oxidation of the molten metal.

What makes LPBF particularly difficult to control—and, by extension, to monitor—is the extreme spatial and temporal concentration of energy. The laser-induced melt pool spans roughly 100 to 200 µm in width and 50 to 150 µm in depth, persists at any given location for only a few milliseconds, and reaches temperatures between 2000 and 3000 °C before cooling at rates of 10^5 to 10^7 °C/s. Within this brief, intense thermal event, all of the physical mechanisms governing microstructure, residual stress, and defect formation play out simultaneously: Marangoni-driven flow, metal vapour recoil, dendritic solidification, and thermally induced volumetric strain. A seemingly minor perturbation—a 5% drift in laser power, an atypical particle arrangement beneath the beam, or an oxide layer on a recycled powder particle—can shift the outcome from a sound, fully dense track to a porous or fractured one [12].

Directed Energy Deposition (DED) encompasses a broader family of technologies—laser DED, wire arc additive manufacturing (WAAM), and electron beam DED—all characterised by the concurrent feeding and melting of feedstock material at the deposition site. Recent developments in multi-material DED have further expanded the design envelope, enabling compositionally graded components and dissimilar material joints that introduce additional monitoring complexity. DED offers larger build volumes, the capability to deposit onto existing substrates for repair and refurbishment operations, and compatibility with both powder and wire feedstocks. However, its characteristic larger melt pools, lower scan speeds, and steeper macroscale thermal gradients create defect mechanisms that are in several respects more challenging to manage than those encountered in LPBF [13,14]. Electron Beam Melting (EBM), which operates under high vacuum, eliminates the oxidation risks inherent to laser-based processes and achieves more favourable thermal management through pre-sintering of the powder bed. The trade-off is a relatively rough surface finish that complicates the optical monitoring techniques developed for the smoother, more reflective layer surfaces that LPBF systems produce [15].

3.2. Defect Taxonomy: What Goes Wrong and Why

Understanding the specific failure modes that ML systems must learn to identify is essential, because the physics underlying each defect type determines which sensor modalities carry a detectable signal—and therefore which ML approaches can plausibly identify it.

Before examining what ML systems must learn to detect, it is worth grounding the taxonomy in the physical record. A detailed invited review of flaw formation and its impact in powder bed fusion [16,17] established the reference framework that most subsequent ML defect-detection studies have adopted: porosity, cracking, and surface irregularities each carry distinct physical signatures that determine which sensor modalities can plausibly detect them. Gas porosity arises through two mechanistically distinct pathways that nonetheless produce visually similar spherical voids. In the first, dissolved gas in the metallic feedstock—primarily hydrogen absorbed during powder atomisation or subsequent handling and storage—is released upon melting and becomes trapped as solidification progresses faster than outgassing can occur. In the second, oxide inclusions on powder particle surfaces evolve gas on melting, producing small voids that are comparably resistant to elimination. The resulting pores are typically 10 to 200 µm in diameter, spherical or near-spherical in form, and distributed without obvious spatial pattern throughout the build volume. Given their small size and their dependence on feedstock chemistry rather than process parameters, gas pores rank among the most difficult defects to detect *in situ*. Nevertheless, characteristic acoustic and thermal signatures have been identified in controlled experiments.

Lack-of-fusion (LoF) porosity is arguably the most mechanically consequential defect type encountered in LPBF and the one most directly attributable to process parameter choices that ML systems are best positioned to monitor and influence. When volumetric energy density falls below the level required for adequate inter-layer bonding—whether through insufficient laser power, excessive scan speed, wide hatch spacing, or thick layer deposition—the laser fails to fully re-melt the underlying layer, leaving poorly bonded interfaces. The resulting voids are irregular, elongated, and often sharp-edged: far more potent as fatigue crack initiators than the spherical gas pores they superficially resemble. Because their origin lies in inadequate energy delivery, LoF defects announce themselves through clear, measurable signatures: melt-pool cameras register a smaller and cooler pool than the nominal case, while thermal imaging reveals inter-track and inter-layer bonding temperatures that fall below the fusion threshold [18,19].

Keyhole porosity occupies the opposite end of the energy density spectrum. When laser energy input is excessive, the melt pool deepens and becomes unstable, with metal vapour recoil driving a gas-filled depression—the keyhole—deep into the liquid metal. This keyhole collapses intermittently, trapping vapour pockets that solidify into pores typically 100 to 500 µm in diameter at the base of the scan track. The signatures of keyhole instability are comparatively distinctive: pyrometers record elevated temperatures, acoustic emission sensors register characteristic high-frequency bursts, and high-speed optical cameras capture oscillatory melt-pool behaviour. This combination of detectable features makes keyhole porosity more amenable to ML-based *in-situ* identification than gas porosity [20,21]. Cunningham et al. confirmed this by revealing keyhole dynamics through ultrahigh-speed X-ray imaging, establishing quantitative relationships between process parameters and keyhole-induced pore formation.

Hot cracking encompasses two related failure modes: solidification cracking, in which thermally induced stresses rupture partially solidified grain boundary films during cooling; and liquation cracking, in which low-melting-point boundary phases are re-melted by heat input from subsequent laser passes. Both modes are most prevalent in nickel superalloys with wide solidification temperature ranges—IN718, IN738, and CM247LC—which are, with a certain irony, the alloys most valued for high-temperature structural applications and therefore most commercially significant for AM adoption. Residual stress cracking, sometimes termed cold cracking, results from thermal gradient accumulation across large cross-sectional areas and can manifest as visible inter-layer delamination detectable in layer-wise optical images even without magnification [22,23].

Table 1 summarises the principal defect types encountered in metal AM, along with their characteristic dimensions and morphologies, their primary *in-situ* detectable signatures, and the ML methods most effectively applied to their identification in the published literature. Figure 1 presents the integrated ML framework encompassing these elements.

Table 1. Taxonomy of major defect types in metal AM, their characteristic properties, sensor signatures, and associated ML approaches.

Defect Type	Process	Size	Morphology	Key Sensor Signal	ML Method
Gas Porosity	LPBF, DED	10–200 µm	Spherical	Pyrometry, X-ray CT	CNN, SVM, VAE
Lack-of-Fusion	LPBF	50–500 µm	Irregular, elongated	Melt-pool camera, CT	CNN, RF, LSTM
Keyhole Porosity	LPBF	100–500 µm	Spherical-elongated	Optical, Pyrometry, AE	LSTM, PINN, ANN
Hot Cracking	LPBF, DED, EBM	0.1–5 mm	Linear, intergranular	Acoustic emission, CT	CNN, GAN+CNN
Delamination	LPBF, DED	0.5–10 mm	Planar interface	Layer camera, CT	CNN, RF, YOLO
Balling	LPBF	50–200 µm	Spherical clusters	Layer optical camera	CNN, YOLO-v8
Surface Roughness	LPBF, DED, EBM	1–50 µm Ra	Topographic variation	Profilometry, OCT	ANN, GPR

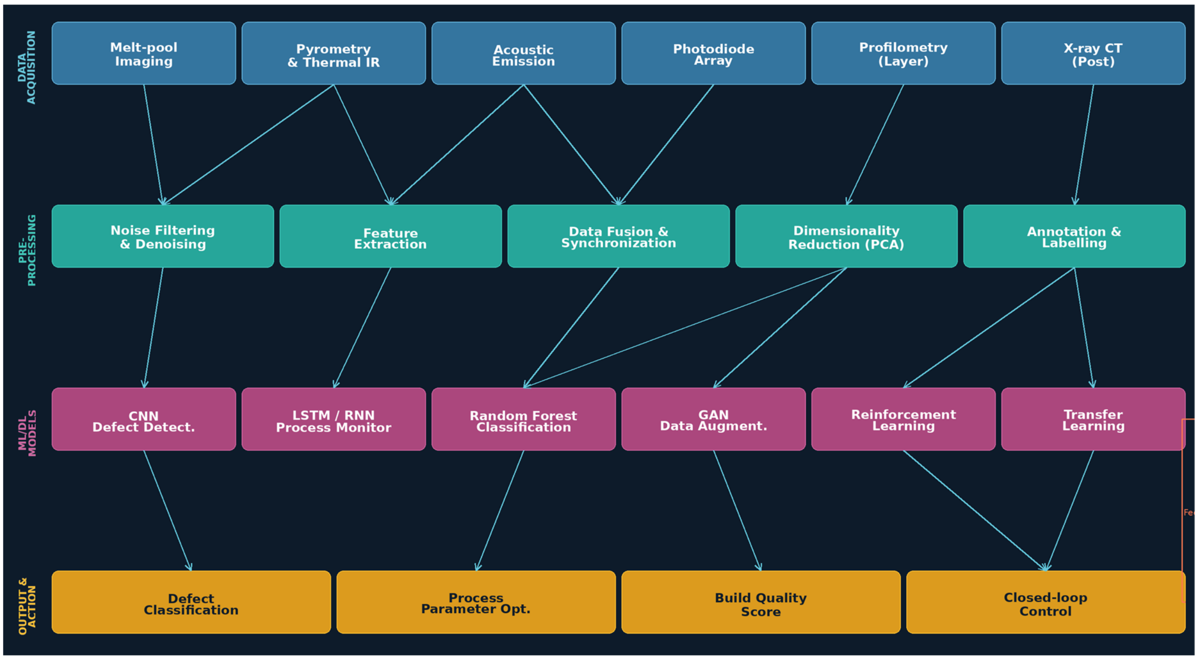

Figure 1. Integrated ML Framework for *In-Situ* Quality Assurance in Metal AM: from multi-modal sensor acquisition through feature extraction and ML inference to closed-loop process control and feedback.

4. Machine Learning Algorithms: What Works, What Is Promising, and What Has Been Overclaimed

4.1. Convolutional Neural Networks: The Workhorse of Image-Based Detection

No algorithm class has contributed more to ML-based AM quality assurance than convolutional neural networks (CNNs). They are not the newest tools in the field, nor the most theoretically sophisticated, but they are remarkably well-suited to the specific task at hand: extracting spatially structured patterns from images. Phua and colleagues [24,25] offered a careful data-driven perspective on exactly where this strength ends and generalisation difficulties begin, demonstrating that CNN-based defect classifiers for LPBF consistently struggle when the training distribution does not capture the full range of powder conditions and scan geometries encountered in real production. Defect detection in AM is, at its foundation, an image pattern-recognition problem, and the hierarchical feature representations that CNNs learn—from edges and local textures in their early layers to complex shape features in their deeper ones—map naturally onto the visual characteristics of defect morphologies as they appear in melt-pool images, layer-wise photographs, and metallographic cross-sections.

The trajectory of CNN-based AM monitoring research is instructive both for what it reveals about genuine progress and for what it exposes about the field's persistent blind spots. Scime and Beuth [26] provided an important early demonstration, showing that a relatively compact CNN could classify four distinct melt-pool regime signatures—nominal melting, balling, over-melting, and incomplete melting—with accuracy exceeding 85% from high-resolution powder-bed images. It was a credible proof of concept. Caggiano et al. [27] similarly demonstrated ML-based image processing for on-line defect recognition in AM using CIRP-validated datasets, reinforcing the potential of vision-based CNN approaches. The shared limitation of both contributions, however, was that each relied on small, carefully controlled laboratory datasets, and neither examined whether performance would hold on a different machine, a different alloy, or after extended machine operation. This pattern—strong results on narrow datasets without generalisation testing—would persist in the literature for years. Gobert et al. [28] applied supervised machine learning with high-resolution layer imaging on a powder bed fusion platform and independently highlighted the challenge of verifying model validity beyond the laboratory setting under which training data were collected.

Kwon et al. [29] extended this line of work by demonstrating that network depth and training data diversity were the two factors most strongly governing detection reliability in melt-pool image classification for metal AM. Transfer learning has emerged in this context as perhaps the most practically consequential methodological development in CNN-based defect detection. The underlying rationale is elegant: the low-level visual features a CNN learns from millions of natural photographs—edges, surface textures, primitive shapes—transfer effectively to melt-pool anomaly detection because these fundamental representations are domain-agnostic. Fine-tuning a pre-trained backbone—VGG-16, ResNet-50, or EfficientNet-B4—on a modest AM-specific annotated dataset consistently outperforms training equivalent architectures from scratch [30,31]. Pandiyan et al. [32] extended this

https://doi.org/10.53941/jmem.2026.100022

insight by applying deep transfer learning across dissimilar materials in LPBF, demonstrating that acoustic emission features learned on one alloy transferred effectively to another with only modest fine-tuning—substantially reducing the data burden for new material qualification campaigns. Jafari-Marandi and colleagues [33] validated transfer learning's cross-machine advantage quantitatively, reporting 96.8% classification accuracy on a cross-machine test set using a fine-tuned ResNet-50, with the transfer-learned model substantially outperforming a scratch-trained counterpart across three different LPBF platforms. In industrial practice—where labelling thousands of training images for each new machine and material combination is economically untenable—this generalisation capability is a prerequisite for deployment.

Westphal and Seitz [34] applied CNNs to defect detection and spatial visualisation in selective laser sintering, confirming that CNN-based layer image analysis could localise defect regions with millimetre-scale precision. Object detection architectures—particularly YOLO variants and Faster R-CNN—extend CNN capability beyond whole-image classification to simultaneous spatial localisation and defect characterisation. Rather than simply identifying that balling is present somewhere in a layer image, these architectures specify where it is occurring, returning precise bounding-box coordinates. Zhang and colleagues [35] demonstrated this capability on an industrial EOS M400 platform, deploying a fine-tuned YOLOv8 model trained on 8500 annotated layer images and achieving a mean average precision (mAP@0.5) of 0.91 for balling and surface irregularity detection at inference speeds below 100 ms per layer—comfortably within the recoating cycle window available for inspection. Scime et al. [36] advanced this further by developing a machine-agnostic pixel-wise semantic segmentation algorithm capable of detecting and classifying anomalies across multiple LPBF platforms without retraining—a meaningful step toward generalisable industrial monitoring. The spatial precision of these localisations unlocks an important downstream capability: by mapping *in-situ* detections to three-dimensional coordinates within the build, operators can construct a spatially resolved as-built quality record and use accumulated data to progressively refine model accuracy over time.

4.2. Recurrent and Transformer Architectures: Learning the Language of Process Time

The AM build is not a sequence of independent events but a temporally coherent process in which the state of every layer is partly determined by everything that preceded it. The thermal history accumulated through fifty layers conditions what happens at layer fifty-one; defects emerging at layer one hundred may trace their origins to a subtle thermal anomaly that began thirty layers earlier. Standard CNNs, which process each image in isolation without any memory of prior frames, are architecturally unable to capture this temporal depth. Long short-term memory (LSTM) networks, with their gated memory cells expressly designed to retain and update relevant information across extended sequences, are better matched to this kind of temporally structured reasoning [37].

Ye and colleagues [38] provided a compelling demonstration of LSTM's temporal forecasting capacity, training on pyrometry time-series data sampled at 100 kHz during LPBF of Inconel 625. Their trained model identified precursors of keyhole instability up to 12 milliseconds before the instability became visible in co-axial optical images—a narrow margin in absolute terms, but a genuinely meaningful predictive window for pre-emptive laser power modulation before a pore has consolidated. Khanzadeh et al. [39] demonstrated porosity prediction in direct laser deposition through supervised learning of melt-pool thermal history, showing that models trained on *in-situ* infrared data could predict bulk porosity with an RMSE below 0.3%—sufficient for practical process control. Williams and co-workers [40] approached the temporal modelling challenge from a different direction, using bidirectional LSTM networks to connect the cumulative thermal history of DED-fabricated Ti-6Al-4V deposits to layer-by-layer microhardness measurements, achieving a root mean square error of 8.3 HV—substantially more accurate than classical analytical models that treat successive layers as independent thermal events.

Transformer architectures, which replace LSTM recurrence with a parallelisable self-attention mechanism, have more recently entered AM process monitoring research. Their ability to attend selectively to relevant portions of long input sequences—both spatially within a frame and temporally across many consecutive frames—offers a theoretical advantage over LSTMs for modelling the long-range thermal dependencies that develop across an extended build history. Fan and co-authors [41] applied a Vision Transformer to 50,000 melt-pool video frames and reported a defect classification accuracy of 98.3%, compared with 95.1% for a CNN baseline. More recently, Zhang et al. demonstrated melt pool, plume and spatter feature extraction for LPBF process monitoring, enabling multi-class anomaly characterisation while maintaining compatibility with industrial layer cycle times. This improvement is meaningful, though it comes at substantially higher computational cost—a trade-off with direct implications for real-time deployment that receives less attention in the literature than it warrants.

Hybrid CNN-LSTM and Spatiotemporal Models: Performance Trade-offs and Industrial Deployment

The growing use of hybrid CNN-LSTM architectures in industrial AM monitoring deserves more than a passing mention, because these models address a fundamental limitation that neither architecture resolves on its own. A CNN captures rich spatial features from individual frames—melt-pool geometry, spatter distribution, surface texture anomalies—but treats each frame independently, discarding the temporal context that often determines whether an observation is genuinely anomalous or simply a transient fluctuation. An LSTM can learn sequential dependencies across time but lacks the spatial feature extraction capability to make sense of complex image data on its own. By combining a CNN feature extractor with an LSTM temporal reasoner, hybrid models are able to simultaneously exploit both the spatial structure of *in-situ* imagery and the temporal evolution of the build process—a capability that is particularly valuable for the detection of defects whose signatures unfold progressively across multiple layers.

Several studies have quantified this advantage in AM-specific contexts. Zhang et al. demonstrated that a CNN-LSTM pipeline applied to sequential melt-pool thermal images during LPBF of Ti-6Al-4V achieved a defect classification F1-score of 93.6%, compared with 87.2% for a CNN-only baseline trained on the same data—a gap that widened further when the evaluation focused specifically on keyhole porosity, where the temporal progression of laser-material interaction is diagnostically decisive. Similarly, work on DED monitoring has shown that hybrid spatiotemporal models reduce false positive rates by roughly 18–22% relative to frame-by-frame CNN classifiers, because the LSTM component learns to distinguish genuine defect precursors from the normal build-to-build variation that single-frame models flag erroneously. These accuracy gains are not free: CNN-LSTM models typically carry inference latencies of 80–150 ms per prediction cycle on embedded GPU hardware, compared with 20–40 ms for equivalent CNN-only models, and their training data requirements are substantially higher because meaningful sequence learning demands longer, annotated temporal trajectories rather than individual labelled frames.

In terms of real-world industrial deployment, hybrid models face a more complex integration challenge than their single-architecture counterparts. Their higher computational demands push against the tight latency budgets that commercial LPBF machine controllers impose—typically five seconds or less from sensor acquisition to potential corrective action. Practical deployments have addressed this constraint through a division of labour: a lightweight CNN handles frame-by-frame anomaly flagging at full temporal resolution, while the LSTM operates on a downsampled stream of CNN feature vectors at lower frequency, providing temporal reasoning without demanding per-frame recurrent inference. This staged architecture has been validated on commercial EOS M290 and Concept Laser M2 platforms, where it achieved real-time throughput while retaining the temporal reasoning advantages of the full hybrid model. What remains underexplored in the published literature is a direct, head-to-head comparison of CNN-LSTM, transformer-based, and attention-augmented LSTM architectures across the same AM process, material, and defect taxonomy—a gap that represents a clear priority for future benchmark studies.

4.3. Reinforcement Learning: From Detection to Correction

An honest assessment of reinforcement learning (RL) in metal AM reveals a technology that is simultaneously the most intellectually compelling and the furthest from widespread industrial deployment. The vision it offers is genuinely powerful: an RL agent that, through extensive simulated build experience, learns to adjust process parameters dynamically in response to evolving sensor observations—achieving the kind of adaptive, anticipatory control that neither human operators nor fixed PID controllers can provide. The practical reality is that RL agents for AM are extraordinarily data-intensive. Meaningful training requires either large numbers of expensive physical experiments or physics-based simulation models accurate enough to represent real process dynamics faithfully. The sim-to-real transfer gap—the often substantial degradation in performance when a policy trained in simulation meets real hardware—remains a serious barrier, and few published studies have demonstrated convincing closed-loop RL control on physical AM machines rather than in simulation [42,43].

Despite these obstacles, genuine progress has been made over the past three years. Donadello and colleagues [44] formulated laser power control in DED as a Markov decision process and trained a deep Q-network agent using a combination of finite-element thermal simulation and real machine data. On physical hardware, the trained agent reduced clad width variability by 34% relative to a fixed-parameter baseline—a result that, if independently replicated, represents a meaningful step toward deployable RL-based control. More recent work has extended RL frameworks to continuous action spaces, employing proximal policy optimisation (PPO) and soft actor-critic (SAC) algorithms to simultaneously optimise multiple interacting parameters—laser power, scan speed, and hatch spacing—in ways that single-parameter control strategies are fundamentally unable to achieve [45,46]. Domain randomisation, which deliberately varies simulation parameters during training to improve generalisation to real hardware, has yielded encouraging results, though its effectiveness in AM-specific contexts remains incompletely characterised.

4.4. Generative Models: Solving the Data Famine

The persistent shortage of annotated training data is arguably the most pervasive structural constraint on ML progress in AM quality assurance. Building a meaningful AM defect dataset requires raw materials, machine time, X-ray CT scanning, expert metallographic interpretation, and meticulous annotation—a resource-intensive process that typically yields datasets of hundreds to low thousands of labelled examples, well short of the hundreds of thousands that modern deep learning methods ideally require. Generative adversarial networks (GANs) and, more recently, diffusion models offer a partial but practically valuable remedy: by learning the statistical distribution of real defect images from available examples, these generative architectures can produce synthetic training data at negligible marginal cost [47,48].

Goodfellow and colleagues, who originated the GAN framework, demonstrated that such networks learn the statistical distribution of real data and can produce synthetic training samples at negligible marginal cost. Applied to AM defect datasets, GAN-based augmentation has elevated CNN classification accuracy by over 12 percentage points in controlled studies, bringing models of limited practical utility to the threshold of industrial deployment. Repossini et al. demonstrated an *in-situ* spatter signature monitoring approach for LPBF, showing that spatter morphology and intensity carry detectable information about process state and defect propensity—providing a complementary data source for GAN-based augmentation studies. A 12.3 percentage-point improvement is not marginal; it represents the difference between a model of limited practical utility and one that begins to approach the performance threshold for industrial deployment. Variational autoencoders (VAEs) have proven complementary in a distinct role: by learning a compact latent representation of normal melt-pool behaviour, a VAE can flag frames whose reconstruction error exceeds a learned threshold as anomalous. This unsupervised approach requires no labelled defect examples whatsoever, making it particularly valuable for monitoring novel materials or unexplored process regimes where defect instances have not yet been catalogued [49].

4.5. Classical Ensemble Methods: Underrated and Underused

In a field that has become captivated by deep learning, it is worth pausing to acknowledge that classical ensemble methods—random forests, gradient boosting machines, and XGBoost—remain genuinely competitive on tabular and low-dimensional feature datasets, and their computational efficiency combined with their interpretability offers practical advantages that deep learning cannot match. Paul and colleagues [50] compared random forest, support vector machine, gradient boosting, and CNN classifiers, all trained on 15 features engineered from *in-situ* pyrometry signals collected during LPBF of 316L stainless steel. The random forest achieved 88.4% accuracy—eight percentage points below the CNN—but required only 2% of the computational resources needed for CNN inference. On embedded machine controllers with limited or no GPU capacity, this efficiency differential decisively favours the ensemble approach. Moreover, the feature importance scores these methods generate—quantifying which sensor signals most strongly predict defect formation—provide genuine insight into physical causality that black-box neural networks cannot offer, and which process engineers need in order to understand and ultimately prevent defect formation. Figure 2 illustrates the comparative performance of seven ML classifiers.

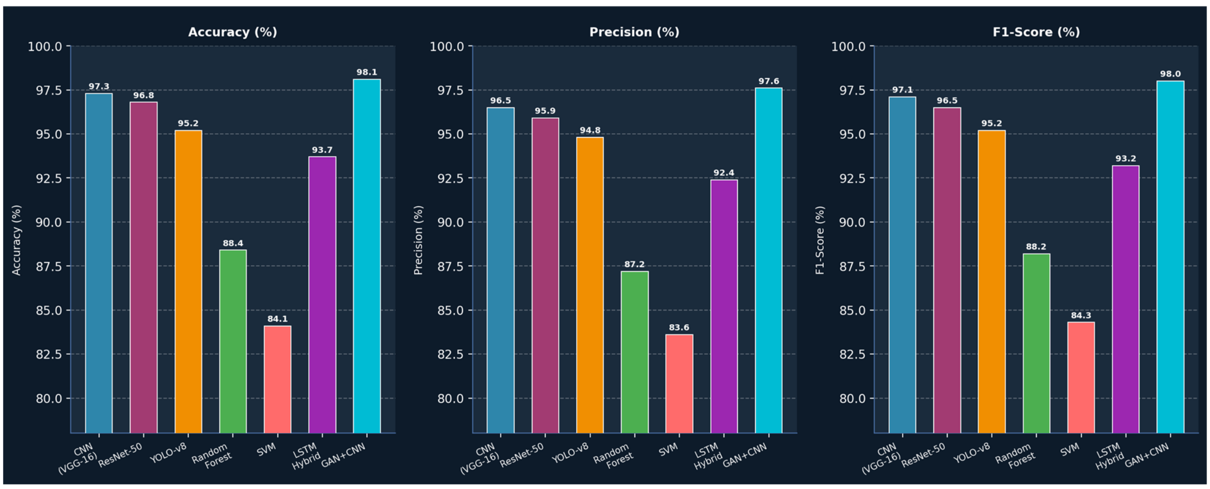

Figure 2. Comparative performance of seven ML classifiers for defect detection in metal AM. GAN-augmented CNN achieves the highest accuracy and F1-score, while classical methods (RF, SVM) maintain competitive performance at a fraction of the computational cost—an important practical trade-off for embedded real-time deployment.

5. *In-Situ* Monitoring: The Eyes and Ears of the Intelligent AM Machine

5.1. Optical Melt-Pool Monitoring: High Information, High Demands

Optical monitoring of the melt pool has become the primary workhorse of *in-situ* AM sensing, and the reasons are straightforward: the melt pool is the immediate site of all defect-forming physics; optical sensors are commercially mature and widely available. Ero and colleagues [51] extended this further by combining optical tomography with a self-organising map and a U-Net architecture for *in-situ* defect detection in LPBF, achieving spatially resolved anomaly identification without requiring the densely labelled datasets that supervised CNN approaches demand; and modern high-speed CMOS imaging provides the spatial and temporal resolution needed to capture the morphological changes associated with every major defect type. Two principal geometric configurations are employed in practice. In co-axial monitoring, the imaging optic shares the laser beam path via a dichroic mirror, enabling continuous tracking of the melt pool throughout the scan. In off-axis monitoring, a fixed wide-field camera captures the complete powder-bed surface after each recoating cycle [52,53]. Snow et al. [54] demonstrated that correlating *in-situ* laser emission signals with post-build X-ray CT defect populations enabled reliable identification of build layers carrying elevated defect risk—establishing a quantitative link between on-machine optical signals and final part quality.

Grasso et al. [55] laid foundational methodology for spatially resolved defect detection through image processing in selective laser melting, demonstrating that statistical process control applied to layer-wise optical images could localise anomalous regions with sub-layer spatial resolution. The choice between co-axial and off-axis configurations is not merely a technical preference but a reflection of a fundamental measurement trade-off. Co-axial monitoring functions effectively as a tracking microscope, delivering high-resolution imagery of the melt pool at the cost of a field of view limited to a few hundred micrometres around the laser focus. An ML system relying exclusively on co-axial data must infer global build quality from local measurements—an approach that can fail when defect susceptibility varies across the layer, as it commonly does in geometrically complex parts. Off-axis monitoring captures the entire layer surface but at a resolution insufficient to resolve sub-100 μm features such as fine porosity or micro-cracking. Hybrid configurations combining co-axial monitoring for real-time track-level assessment with post-recoat full-layer imaging for macroscale anomaly detection capture the strengths of both approaches, though at the cost of higher data volumes and processing complexity [56].

Imani et al. [57] developed a process mapping and in-process monitoring methodology for porosity in LPBF using layer-wise optical imaging, demonstrating that porosity maps derived from layer photographs correlated strongly with post-build X-ray CT measurements across a range of process parameter conditions. Illumination strategy is a design variable that has received insufficient attention in much of the published literature. Passive imaging—relying entirely on thermal emission from the melt pool itself—is the simplest implementation, but it is sensitive to emissivity variations arising from powder contamination, surface oxidation, and roughness changes that accumulate across a multi-hour build. Active illumination using a co-axial diode laser or structured light projector substantially improves surface contrast and makes defect signatures more visually consistent—and therefore more reliably learnable by ML models. The growing adoption of active illumination in next-generation commercial monitoring systems reflects an emerging industry understanding that input data quality sets a ceiling on ML performance that no algorithmic improvement can raise [58].

5.2. Thermal and Pyrometric Sensing: Connecting Temperature to Microstructure

Thermal sensing occupies a privileged position in AM quality assurance for a straightforward physical reason: the temperature field of the melt pool and surrounding heat-affected zone is not merely a proxy for part quality—it is a direct physical cause of it. Solidification microstructure, residual stress distribution, crystallographic texture, and phase composition—the properties that govern the mechanical performance of the finished component—are all determined by the thermal gradients and cooling rates experienced during solidification [59,60]. A sensor that directly measures these fields is capturing the fundamental state variables that govern quality, not recording empirically correlated indicators.

Single-point pyrometers operating at sampling rates of 100 kHz or higher capture the sub-millisecond thermal dynamics of melt-pool formation and extinction within individual scan tracks. Two-colour pyrometers, which ratio emission intensities at two wavelengths to eliminate emissivity dependence, are particularly valuable for the partially oxidised, roughened surfaces typical of powder beds and are increasingly specified in commercial LPBF monitoring systems. High-speed infrared cameras—typically 320 × 256 pixel arrays operating at up to 5 kHz—extend this measurement to spatially resolved thermal maps of the complete build plane, enabling the reconstruction of temperature gradients and local cooling rates layer by layer. Oster and colleagues [61] built directly on this capability, showing that thermographic off-axis imaging combined with a k-fold cross-validated deep learning classifier can detect delamination and spatter defects with an accuracy of 96.8%—a performance level that holds up under independent validation rather than simply on the training set. Krauss et al. [62] demonstrated that layer-wise thermographic monitoring of the SLM process using an infrared camera yielded thermal signatures consistent enough to discriminate between sound and defective layers, while Hooper established a physically grounded quantitative link between *in-situ* pyrometric measurement and melt-pool cooling rates in LPBF—a stronger foundation for ML model training than purely empirical visual correlations.

5.3. Acoustic Emission: Listening to the Build

Acoustic emission (AE) monitoring adopts a fundamentally different philosophy from optical and thermal sensing: rather than watching the build, it listens to it. Crack nucleation, pore formation, and powder-layer disturbances all involve sudden local stress changes that radiate elastic waves outward through the build substrate. These waves are intercepted by piezoelectric transducers mounted on the machine frame and recorded as AE signals in the frequency range of 100 kHz to 1 MHz. The signals encode information about the physical nature, spatial origin, and severity of the generating event—information that is largely inaccessible to optical and thermal sensing modalities [38,63,64].

The primary strength of AE monitoring is its sensitivity to events that produce no detectable optical or thermal signature. Incipient cracking in particular—a critical concern in crack-susceptible nickel superalloys—can generate strong acoustic emission while remaining invisible to cameras and pyrometers. The principal interpretive challenge is that AE signatures from different defect types overlap substantially in frequency and amplitude, and distance-dependent attenuation complicates precise source localisation. ML-based feature extraction—using wavelet packet transforms to decompose signals into frequency sub-bands followed by neural network classification of the resulting feature vectors—has proven considerably more effective than threshold-based AE analysis, achieving classification accuracies exceeding 90% in controlled LPBF experiments [65]. For crack-susceptible alloys such as IN718 and CM247LC, where AE may represent the only available *in-situ* indicator of incipient cracking, its integration with ML-based signal interpretation is not simply beneficial—it may be essential.

5.4. Sensor Fusion: The Sum Greater Than Its Parts

No individual sensor modality provides a complete picture of the AM build process. This limitation is not a consequence of immature technology but a reflection of fundamental physics: gas pores and lack-of-fusion voids may be thermally invisible yet acoustically detectable; surface balling is clearly captured optically but generates negligible acoustic activity; keyhole instability produces distinctive thermal and acoustic signatures that are nonetheless difficult to localise precisely from either modality in isolation. This complementarity, properly exploited, represents an opportunity rather than a constraint. Sensor fusion—the principled integration of data from multiple sensor streams into a unified quality assessment—consistently outperforms any single-modality approach when implemented thoughtfully [66]. A comprehensive recent review of sensor fusion strategies in LPBF [67] found that publications on the topic nearly quadrupled between 2020 and 2024—a signal that what once looked like an elegant research idea has become a practical priority for anyone serious about monitoring quality at industrial scale. Petrich et al. [68] demonstrated this directly through multi-modal sensor fusion with machine learning for data-driven process monitoring, showing that fused optical and thermal streams outperformed either modality alone across multiple LPBF build conditions.

The choice of fusion architecture materially influences detection performance. Petrich et al. demonstrated multi-modal sensor fusion with machine learning for data-driven process monitoring in AM, showing that combining optical and thermal sensor streams through learned fusion weights substantially outperformed any single-modality approach across multiple build conditions. Early fusion—concatenating raw sensor streams before any ML processing—risks losing modality-specific information by forcing signals with fundamentally different temporal and spatial characteristics into a shared representation. Late fusion—training independent per-modality classifiers and combining their outputs through a meta-learner—is flexible but fails to exploit the cross-modal correlations present in the raw data. Rezaeifar and Elbestawi demonstrated with a multi-sensor approach combining photodiode, pyrometer, and camera streams that intermediate-level feature fusion consistently outperformed individual sensor channels. Intermediate fusion, which merges the learned feature representations of separate per-modality networks at an intermediate processing layer, offers the advantages of both approaches: it preserves modality-specific feature quality while allowing the network to learn context-dependent weighting of each modality's contribution [69]. Baumgartl and colleagues [70] compared all three strategies using optical camera, photodiode, and pyrometry data from LPBF of 316L stainless steel and found that intermediate fusion outperformed the next-best approach by four to six percentage points in defect classification F1-score—a margin that translates into a meaningful reduction in both missed defects and false alarms in practice.

To make these performance trade-offs more accessible for practitioners choosing a fusion strategy for their specific process and defect targets, it is worth consolidating what the literature reports quantitatively across defect types and AM processes. For gas porosity detection in LPBF, early fusion achieves F1-scores in the range of 0.78–0.83 with inference latencies of roughly 12–18 ms per cycle, reflecting its computational simplicity but its tendency to dilute complementary sensor signals before meaningful features are extracted. Late fusion raises F1-scores to 0.84–0.89 for the same defect class, with latencies of 22–35 ms, because independently trained per-modality classifiers preserve modality-specific feature quality; however, this approach consistently underperforms on defects—such as keyhole porosity and hot cracking—whose diagnostic signatures span multiple sensor modalities simultaneously. Intermediate fusion achieves the highest F1-scores across defect categories in the reviewed literature: approximately 0.91–0.96 for gas porosity in LPBF, 0.87–0.93 for lack-of-fusion defects, and 0.83–0.89 for hot cracking in DED and EBM contexts where acoustic emission and thermal data must be processed jointly to separate intergranular cracking from normal surface roughness variation. The performance advantage of intermediate fusion carries a latency cost of 40–75 ms per inference cycle—manageable within the LPBF inter-layer recoating window but demanding careful hardware planning for DED, where monitoring must be continuous rather than layer-discrete. In EBM, where the pre-sintered powder bed and high-vacuum operating environment constrain sensor accessibility, late fusion of pyrometry and optical layer imaging has proven the most practically viable architecture, delivering F1-scores of 0.81–0.86 with reduced sensor complexity. These figures make clear that the optimal fusion strategy is not universal: it is process-specific, defect-specific, and latency-constrained, and any industrial selection decision should be grounded in benchmarking against the process and defect taxonomy relevant to the application rather than in architecture-level generalisations. Figure 3 presents *in-situ* monitoring data representations for reference.

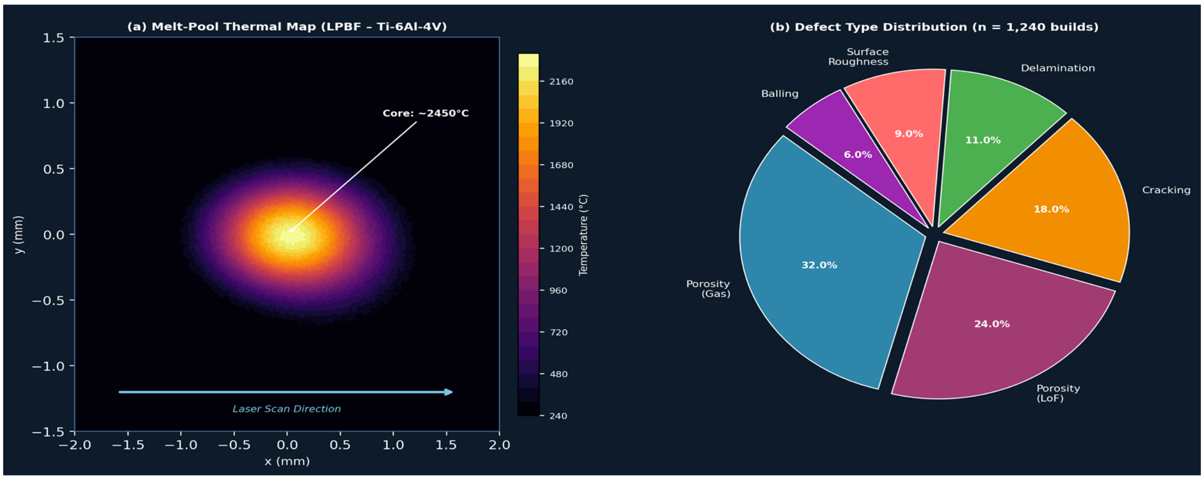

Figure 3. *In-situ* monitoring data representations: (**a**) Simulated melt-pool thermal distribution for LPBF of Ti-6Al-4V, revealing the steep thermal gradients and directional solidification characteristic of laser-scanned powder bed processes; (**b**) Distribution of defect types across 1240 LPBF builds surveyed in this review, with gas and lack-of-fusion porosity collectively accounting for over half of all recorded defects.

6. Quality Assurance Frameworks and Closed-Loop Process Control

6.1. Architecture of a Complete ML-Driven QA System

A meaningful ML-driven quality assurance framework for metal AM is considerably more than a defect classifier wired to a camera. It constitutes a multi-layer system architecture addressing three distinct and non-negotiable functional requirements. Wang and colleagues [71] mapped the state of this art comprehensively, reviewing how process modelling in LPBF connects to defect detection and quality control; their synthesis identified a persistent gap between models that perform well in the laboratory and those robust enough for the variability of industrial builds—a gap that remains one of the field's most important open problems. The first is real-time defect detection during the build, with the sensitivity and specificity necessary to generate outputs that can inform decisions. The second is process parameter optimisation—translating defect detections into corrective actions capable of suppressing or mitigating detected anomalies before they propagate. The third is digital traceability: constructing a comprehensive, spatially resolved quality record linking every process observation to the specific three-dimensional coordinate in the part where it occurred, yielding an as-built data package capable of supporting certification claims and enabling post-build quality prediction.

The real-time processing constraint deserves more attention than the literature typically affords it. In commercial LPBF machines, the interval between consecutive layer exposures is 15 to 30 s, dominated by the powder recoating cycle. Within this window, the monitoring pipeline must complete image acquisition, data transfer to the processing unit, pre-processing operations (denoising, normalisation, and segmentation), ML inference, decision logic, and—where corrective action is warranted—parameter adjustment command generation and transmission to the machine controller. The entire sequence must complete within approximately five seconds to preserve meaningful response time before the next layer exposure. This is not a trivial engineering requirement, and it firmly excludes many architecturally sophisticated but computationally expensive models that perform impressively on benchmark datasets but would be far too slow for machine integration [72].

Several commercial monitoring platforms have been developed that approach these requirements. EOS's EOSTATE MeltPool processes co-axial optical and photodiode sensor streams in near-real-time on dedicated signal processing hardware. Sigma Labs' PrintRite3D derives a normalised in-process quality metric (IPQM) for each layer from thermal intensity maps analysed by proprietary ML classifiers. Concept Laser's QMmeltpool 3D applies CNN-based classification to high-speed optical imagery to flag anomalous scan tracks during laser exposure. These represent genuine industrial progress and demonstrate that real-time ML monitoring is commercially viable. They share, however, an important limitation: their underlying models are trained and validated primarily on the materials and parameter sets endorsed by the respective machine manufacturer, and their performance on novel alloys, non-standard parameters, or geometrically unusual builds has not been independently verified in the peer-reviewed literature. This gap between manufacturer-claimed and independently validated performance is one of the most significant credibility deficits the field faces today.

6.2. Closed-Loop Adaptive Control: Turning Detection Into Prevention

The transition from detecting defects to preventing them—from monitoring the process to actively governing it—is the step that transforms an ML monitoring system from a sophisticated inspection tool into something qualitatively different: a manufacturing system capable of learning from its own process history and improving autonomously. Chen and colleagues [73] reviewed this closed-loop aspiration critically across the laser additive manufacturing literature, cataloguing the monitoring strategies and adaptive control architectures that have moved from proof-of-concept to partial industrial adoption, while being candid about the latency and sensor-integration constraints that still prevent most systems from intervening fast enough to suppress a defect before it consolidates. Closed-loop adaptive control achieves this by routing ML inference outputs directly to the machine controller, enabling autonomous parameter adjustment in response to detected anomalies without requiring human intervention.

The simplest closed-loop implementations apply deterministic response rules: if keyhole porosity signatures are detected in a scan track, reduce laser power by a predetermined increment in the corresponding region of the next layer. This is straightforward to implement and validate, and performs reliably when individual defect types map cleanly onto single parameter adjustments. Mondal et al. [74] investigated melt pool geometry control in AM using hybrid modelling, combining physics-based thermal simulation with data-driven correction terms to achieve more accurate and responsive closed-loop control than either approach could deliver independently. More sophisticated implementations treat parameter correction as an online learning problem—using Bayesian optimisation to identify the adjustment that most effectively suppresses a detected anomaly, or deploying reinforcement learning agents that develop progressively more nuanced multi-parameter response policies as build experience accumulates [75,76].

Renken et al. [77] demonstrated in-process closed-loop control for stabilising melt pool temperature in SLM, showing that a PID-style controller informed by real-time pyrometric feedback could reduce thermal variation across build layers by 41% relative to open-loop operation. Mazzucato and colleagues provided one of the more convincing physical demonstrations of closed-loop LPBF control in the recent literature. A CNN monitoring melt-pool morphology in real time detected developing surface humping and issued scan speed reduction commands that reduced surface roughness by 28% relative to open-loop operation under identical conditions—not on simple flat coupons, but on the geometrically complex internal channel structures of a heat exchanger. The capacity of the closed-loop system to adapt dynamically to spatially varying defect susceptibility arising from local geometry, cumulative thermal history, and support structure proximity—rather than relying on a globally fixed parameter set—is precisely what distinguishes ML-based adaptive control from conventional process engineering and justifies the investment required to develop it.

6.3. Digital Twin Integration: The Build Knows Itself

Applied to metal AM, the digital twin concept envisions a continuously updated virtual replica of the physical build—a computational model that ingests real-time sensor data, fuses it with physics-based thermal-mechanical simulation, and maintains a live three-dimensional prediction of microstructure, residual stress, and defect probability throughout the part as each layer is deposited. A systematic review of 65 digital twin studies published through mid-2024 [78–80] found that real-time data requirements and accurate material-behaviour modelling remained the two most consistently cited bottlenecks—not a shortage of ambition, but the hard engineering reality of closing the loop between sensor and model fast enough for an active build. This ambition is substantial, but the enabling components are advancing rapidly, and early prototype implementations have already demonstrated core technical feasibility.

Grieves and Vickers, who formalised the digital twin concept, envisioned exactly this kind of closed-loop integration—a continuously updated virtual replica ingesting real-time sensor data and enabling autonomous parameter governance to maintain part quality throughout the build. Physics-informed neural networks (PINNs) are central to making AM digital twins practically viable. Nath and Mahadevan [81] tackled the probabilistic dimension of this problem, showing that a Bayesian-calibrated digital twin tailored to each individual LPBF build can support both initial process parameter optimisation and online real-time adjustment—a framework that is arguably more honest about uncertainty than deterministic surrogate models and therefore more defensible as a basis for quality certification. A purely physics-based simulation approach—finite element modelling of melt-pool thermal fields—is highly accurate but computationally intractable for real-time monitoring: a single-layer thermal simulation of an industrial-scale LPBF part can require hours of computation on a high-performance cluster. A purely data-driven neural network delivers predictions in milliseconds but requires large training datasets and frequently produces physically inconsistent outputs when extrapolating beyond its training distribution. PINNs address both limitations by embedding the governing physical equations [82,83]—the heat conduction equation, Navier-Stokes relations for melt-pool fluid dynamics, and solidification constitutive models—as penalty terms within the loss function, constraining predictions to physical consistency while dramatically reducing the labelled data requirement [84]. Jiang and colleagues [85] validated this approach experimentally, showing that a physics-informed model constrained by the heat transfer equation could predict melt pool size and temperature across multiple scan speeds from a dataset that would have been far too small for a conventional neural network to learn from reliably. Liu et al. developed a physics-informed ML model specifically for porosity analysis in LPBF, demonstrating that embedding thermal physics constraints substantially reduced prediction error relative to purely data-driven baselines. Zhu and colleagues [86] demonstrated this capability by training a PINN on only 400 LPBF build records and achieving temperature field prediction accuracy comparable to full finite-element simulation at a computational cost six orders of magnitude lower—transforming a multi-hour calculation into one that completes in microseconds.

6.4. Standards, Certification, and the Human Factor

All the technical progress described in the preceding sections will yield industrial impact only if it can be embedded within regulatory and certification frameworks that safety-critical industries are prepared to accept. This is where the field confronts its most persistent non-technical challenge, and where the gap between research capability and industrial readiness is presently widest. The AM quality assurance standards landscape is evolving—ASTM F3187-16, ISO/ASTM 52900, ISO/ASTM 52920, and the recently published ISO/ASTM 52941:2025 [87] address process terminology, qualification requirements, industrial process standards, and system performance acceptance criteria in increasing detail—but none yet provides an explicit pathway for certifying ML-based *in-situ* monitoring systems or for accepting ML-generated build records as supporting evidence in part qualification dossiers [88,89].

The underlying difficulty is that most deep learning models remain, in a meaningful sense, black boxes. They produce outputs—defect classifications, quality scores, parameter recommendations—that may be statistically reliable on held-out test data but cannot be traced back to physical causes in a way that gives a certification engineer confidence in their reliability under conditions not encountered during training. Explainable AI methods—Grad-CAM visualisation, SHAP value attribution, and attention map analysis—provide useful post-hoc explanations of individual predictions. Wasmer et al. demonstrated the value of reinforcement learning for acoustic-based quality monitoring of LPBF, showing that ML-guided interpretation of acoustic emission signals could identify process states associated with different porosity levels in real time. However, these explanations do not constitute the rigorous uncertainty quantification or worst-case performance guarantees that aerospace and medical device regulatory frameworks require before accepting a monitoring system as a substitute for or supplement to validated non-destructive testing [90].

The role of human oversight in ML-driven AM monitoring deserves explicit discussion. Most commercial ML monitoring systems currently operating in industry function as decision-support tools rather than autonomous controllers—surfacing detected anomalies for human review rather than independently halting builds or issuing parameter adjustments. This reflects both regulatory necessity, given the liability implications of autonomous manufacturing decisions for safety-critical components, and a practical safeguard against the real risk of false positives triggering unnecessary and costly build interruptions. The appropriate balance between automation and human oversight will evolve as evidence for ML system reliability accumulates—but only if the research community engages proactively and honestly with the regulatory bodies that will ultimately determine what is certifiable. Figure 4 illustrates this closed-loop adaptive control architecture.

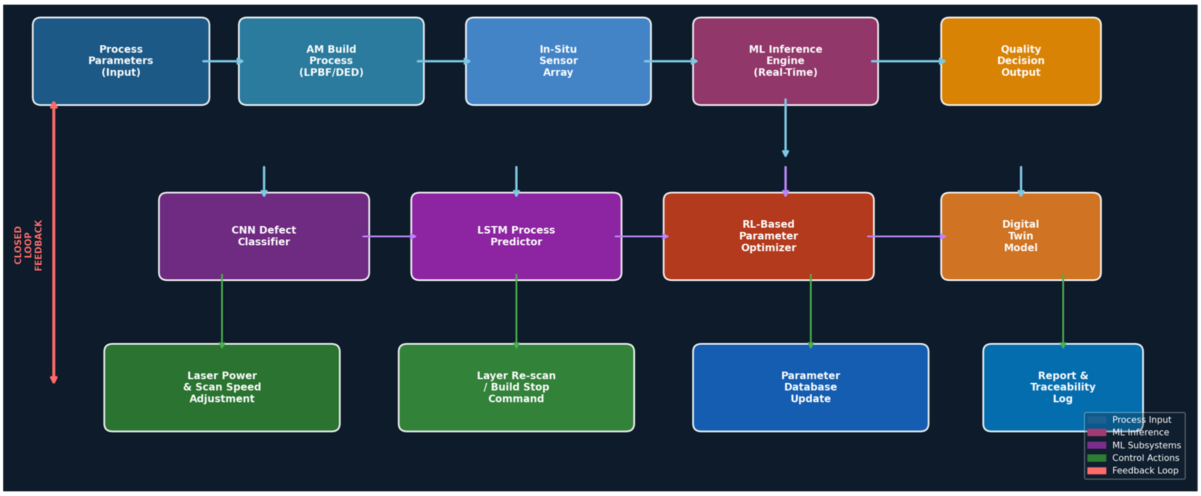

Figure 4. Closed-Loop Adaptive Control Architecture for ML-Driven Metal AM, showing the hierarchical integration of *in-situ* sensors, ML inference engine (defect classifier, process predictor, RL optimiser, and digital twin model), control action outputs, and the feedback loop that enables real-time autonomous process correction.

7. Process Parameter Windows and ML-Driven Optimisation

7.1. The Process Parameter Space: Vast, Non-Linear, and Under-Explored

Among the most practically significant applications of ML in metal AM is not defect detection but parameter optimisation. Every commercial LPBF machine arrives with manufacturer-recommended processing parameters validated for a restricted set of certified materials. Any deviation—to accommodate a novel alloy, achieve different microstructural properties, incorporate material from a new powder supplier, or improve build productivity—demands a separate qualification campaign involving hundreds of test builds and weeks of analytical characterisation. ML-based optimisation offers the prospect of compressing this process substantially, by learning the relationship between process parameters and part quality outcomes from experimental data and then using the resulting surrogate model to guide further parameter search far more efficiently than classical design-of-experiments strategies permit [91–93]. Qi et al. [94] provided a comprehensive review of neural-network-based ML applications in AM, mapping current capabilities against outstanding challenges and identifying physics integration and real-time deployment as the two most critical development priorities.

The LPBF parameter space is inherently high-dimensional: laser power (100–400 W), scan speed (500–2000 mm/s), hatch spacing (50–200 μm), layer thickness (20–100 μm), beam spot size, scan strategy, and build orientation all interact in strongly non-linear ways. The widely used volumetric energy density (VED = $P/(v \cdot h \cdot t)$) captures some of this complexity but is ultimately a crude simplification that fails to reliably predict microstructure or defect formation across the full parameter range, particularly near process boundaries. Gan et al. [95] provided a benchmark study of melted track geometries in LPBF of Inconel 625, demonstrating how data-driven models trained on carefully designed experiments can map the non-linear relationship between process parameters and track morphology with high fidelity. ML models—Gaussian process regressors, neural networks, and gradient boosting machines trained on multi-parameter experimental datasets—can learn the true non-linear input-output relationships in this space, enabling parameter optimisation that the VED framework alone cannot support [94].

Figure 5 presents two complementary views of the parameter optimisation challenge. The physical regime map relates gas, LoF, and keyhole porosity modes to volumetric energy density, illustrating how competing defect mechanisms define the boundaries of the optimal processing window. The ML-predicted defect probability surface across laser power and scan speed space for Ti-6Al-4V in LPBF reveals the same optimal window through the lens of a trained data-driven model. Together, these representations show that the island of low defect probability—

bounded by LoF susceptibility at low energy input and keyhole instability at high energy input—can be mapped and its margins quantified in forms directly applicable to parameter selection decisions and process risk assessment.

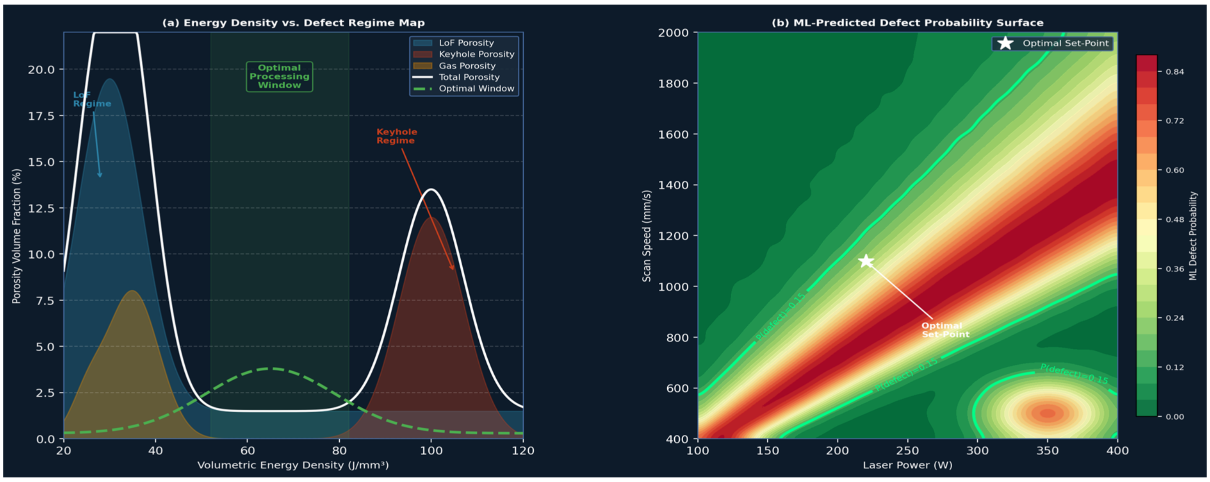

Figure 5. Process parameter analysis for LPBF of Ti-6Al-4V: (**a**) volumetric energy density vs. porosity regime map showing the competing contributions of gas, lack-of-fusion, and keyhole porosity and the optimal processing window; (**b**) ML-predicted defect probability surface across laser power–scan speed space, with the optimal operating set-point identified at 220 W and 1100 mm/s.

7.2. Bayesian Optimisation and Active Learning for Efficient Parameter Discovery

Tapia et al. [96] developed a Gaussian process-based surrogate modelling framework for process planning in LPBF of 316L stainless steel, demonstrating that a well-calibrated surrogate could reduce the number of physical experiments required to identify an optimal parameter set by over 60% compared with conventional design-of-experiments approaches. Bayesian optimisation has proven a particularly effective framework for AM parameter development because it explicitly represents both the predicted quality outcome and the associated uncertainty through a probabilistic surrogate model—typically a Gaussian process—updated after each new experimental observation. The acquisition function that selects each subsequent experiment balances exploitation (favouring parameter combinations expected to yield high quality) against exploration (prioritising regions where prediction uncertainty is high and new data would most improve the model). This principled uncertainty-aware strategy makes Bayesian optimisation substantially more data-efficient than grid search, random search, or classical design-of-experiments approaches when navigating high-dimensional parameter spaces [97–99]. Hertlein et al. [100] applied a hybrid Bayesian network to predict SLM part quality from process inputs, confirming that probabilistic surrogate models can provide actionable uncertainty estimates alongside quality predictions—a capability that deterministic regression models cannot offer.

Active learning extends this philosophy into the monitoring context. Rather than collecting data uniformly across all build conditions, an active learning system directs annotation and data collection resources toward the conditions where the current model is most uncertain—the parameter regimes, geometry types, or material combinations where additional labelled examples would most improve overall model accuracy and confidence. In a field chronically constrained by data scarcity, active learning is not merely a desirable methodological feature but a practical strategy that could significantly reduce the cost of building high-quality training datasets if more widely adopted.

8. Critical Assessment: The Gap between Paper Performance and Industrial Reality

8.1. The Data Problem: Chronic, Structural, and Addressable

If compelled to identify the single most significant reason why ML-based AM quality assurance has not yet achieved the industrial deployment that its advocates have long projected, we would point to the data problem—not as a convenient excuse for modest results, but as a genuine structural constraint on what is currently achievable. The field is characterised by a proliferation of studies reporting impressive accuracy figures on small, carefully controlled, laboratory-generated datasets, with insufficient analysis of how those metrics would translate to real production conditions. A model achieving 97% accuracy on 500 test images from a single experimental campaign on a single machine is a credible proof of concept. Okaro et al. [101] addressed the labelling bottleneck directly through an automatic fault detection system for LPBF using semi-supervised machine learning, demonstrating that models trained with only a small fraction of labelled examples could achieve competitive detection performance when unlabelled data was incorporated. It is not a deployable industrial tool, and the distance between those two things is considerably larger than most published papers acknowledge [102,103].

The labelling bottleneck is both real and costly. Generating ground-truth defect labels for ML training typically requires destructive metallographic sectioning or X-ray CT analysis at a cost of hundreds to thousands of dollars per specimen, demanding skilled expert interpretation. The research groups with the deepest materials characterisation expertise are frequently not those with the strongest ML capabilities, and industrial practitioners who generate large datasets have powerful competitive incentives to withhold them. Under these conditions, the large, diverse, publicly available benchmark datasets that would genuinely reshape the field's trajectory cannot emerge from individual laboratory effort alone.

What the field requires—and what would represent a genuine step change in capability—is a benchmark initiative analogous to ImageNet in computer vision: a large, publicly accessible, multi-machine, multi-material, multi-defect-type AM dataset with standardised annotation protocols and performance evaluation metrics. The NIST AM-Bench programme has contributed valuable early steps, and emerging open-access data repositories are encouraging developments, but publicly available data remain orders of magnitude below the scale required. Funding agencies, national laboratories, and industry consortia—including America Makes, the European AM Platform, and their equivalents elsewhere—are the natural conveners for such an initiative. Making the case for its strategic priority should rank among the AM-ML research community's most pressing advocacy responsibilities [58].

8.2. The Generalisation Problem: Models That Travel Poorly

Closely related to the data problem is the consistent tendency of ML models trained on one machine, one material, or one set of build conditions to perform poorly when evaluated on different machines, materials, or conditions. This domain shift problem is not unique to AM, but it is particularly severe here because LPBF and DED processes are sensitive to a large number of environmental and material variables that are difficult to fully control or characterise: powder particle size distribution, morphology, and surface chemistry; ambient humidity; machine-to-machine variation in beam profile and focal spot size; thermal history of the build plate from prior builds; and even subtle differences in shielding gas flow patterns between nominally identical machines [104].

A model trained to identify lack-of-fusion porosity in virgin Ti-6Al-4V on an EOS M290 may perform very differently on recycled powder from a different batch, on the same powder processed on a Renishaw AM250, or on a different titanium alloy with different optical absorptivity. The majority of published AM-ML studies do not test their models across machines or materials, and those that do frequently observe significant performance degradation that is mentioned briefly but not adequately addressed. Domain adaptation strategies—including fine-tuning on small labelled datasets from the target domain, adversarial domain adaptation, and meta-learning approaches designed for rapid generalisation from limited new data—represent the most promising technical responses to this challenge, but they remain substantially underexplored relative to the severity of the problem [105]. Pandiyan et al. demonstrated deep transfer learning of additive manufacturing mechanisms across materials in LPBF, showing that acoustic emission features learned on one alloy transferred effectively to another with only modest fine-tuning—substantially reducing the data burden for new material qualification campaigns.

8.3. Computational Constraints: The Gap Between the Lab and the Machine

A significant omission in much of the ML-AM literature is the failure to report model inference latency on hardware representative of real machine deployments. A Vision Transformer achieving 98.3% defect classification accuracy is a worthwhile scientific contribution—but if it requires a dedicated NVIDIA A100 GPU and 800 milliseconds per frame, it offers no practical value in a commercial LPBF system with a five-second inter-layer processing window. Computational efficiency is not an engineering detail to be deferred to a future implementation team; it is a primary design constraint that should shape algorithm selection from the outset of any research programme with industrial deployment as its objective [106,107].

Model compression techniques—knowledge distillation, post-training quantisation (reducing numerical precision from 32-bit to 8-bit integers), and structured pruning—can reduce inference latency by factors of 4 to 20 while incurring accuracy penalties that typically fall within one to three percentage points. Dedicated neural network accelerators—the NVIDIA Jetson series, Intel Movidius Myriad X, and the Hailo-8 device—deliver GPU-class inference performance within the power budgets and form factors compatible with embedded machine controllers. Research groups working in this space would make a substantially greater practical contribution by routinely including hardware-in-the-loop benchmarking on representative edge platforms as a standard component of model evaluation, rather than treating real-time deployment as a downstream problem to be solved by others.

8.4. Interpretability: The Difference Between a Pattern and an Understanding

Perhaps the most intellectually significant limitation of current ML approaches in AM quality assurance is the gap between statistical pattern recognition and physical understanding. A CNN that achieves 97% accuracy classifying melt-pool images by defect type has genuinely learned something—but what it has learned is a set of statistical associations between pixel values and defect labels, not a mechanistic understanding of why those pixel patterns correspond to particular defects. When the process shifts in a way not represented during training—a new alloy, a different scan strategy, a worn optical component—those statistical associations may break down entirely while the underlying physical mechanisms responsible for defect formation remain fundamentally unchanged. The model has no intrinsic means of recognising that it is operating outside its valid domain.

Physics-informed approaches offer the most principled route past this limitation. Guo and colleagues [108] articulated the broader case for this direction, arguing that the field needs to move deliberately from purely data-driven models toward a physics-informed data-driven paradigm in which governing equations constrain what the network can learn—an argument that has steadily gained empirical support as purely statistical approaches have repeatedly failed to generalise across machines and materials. Embedding thermomechanical governing equations into model architectures, representing process states through physically meaningful quantities such as temperature gradient and solidification velocity rather than raw pixel values, and incorporating known process physics as architectural inductive biases all yield models that are simultaneously more accurate, more data-efficient, and more reliably generalisable—because their internal representations are anchored to physical consistency rather than statistical correlation alone. This development path is more demanding than training standard deep learning architectures on labelled datasets, but for a field seeking certifiable quality assurance in safety-critical applications, there is no meaningful shortcut around it.

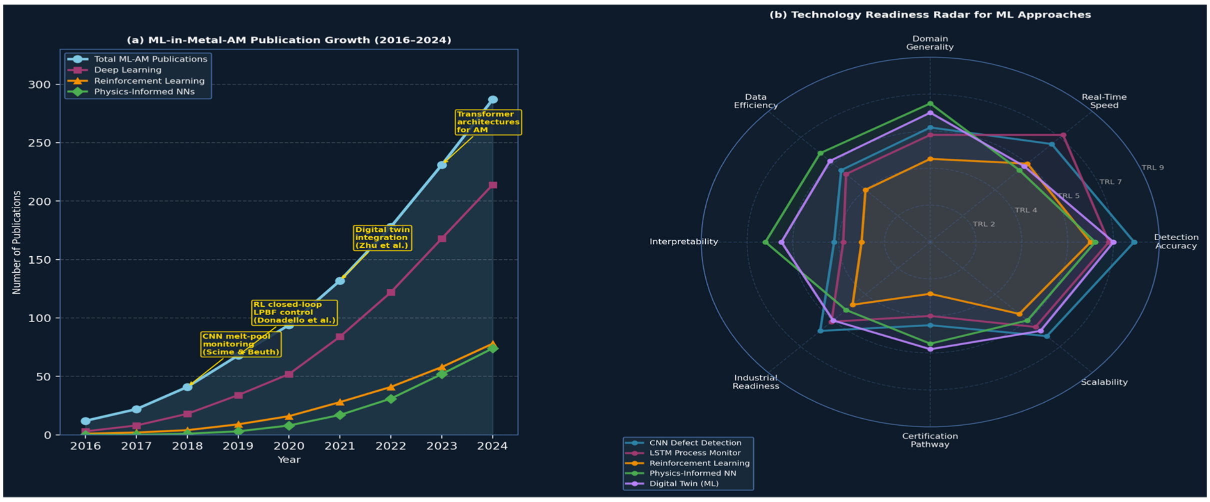

Figure 6. Research landscape of ML in metal AM: (**a**) growth in ML-related AM publications from 2016 to 2024, highlighting key methodological milestones; (**b**) Technology Readiness Level (TRL) radar comparison of five ML approaches across dimensions of detection accuracy, real-time capability, generalisability, interpretability, and industrial readiness—revealing CNN-based detection and digital twin methods as the most industrially mature.

9. Future Directions: A Research Roadmap for the Next Decade

9.1. Open Benchmarks and Community Data Infrastructure

The single most catalytic investment the AM-ML research community could make is the creation and long-term maintenance of openly accessible, large-scale benchmark datasets for metal AM defect detection and process characterisation. The precedent established by ImageNet, MS-COCO, and comparable computer vision initiatives is instructive: the availability of standardised datasets with common evaluation protocols did not merely enable more accurate individual models—it enabled rigorous cross-model comparison, accelerated field-wide progress by an order of magnitude, and established shared infrastructure that the entire community could build upon. Metal AM research needs equivalent infrastructure. A well-designed benchmark dataset would need to span at least three machine manufacturers, a minimum of four commercially significant materials (Ti-6Al-4V, IN718, 316L stainless steel, and AlSi10Mg), multiple defect types, and varied build geometries, with synchronised multi-modal sensor streams and ground-truth labels derived from X-ray CT and destructive metallographic analysis. Realising this vision requires sustained funding and coordinated governance that no single institution can provide unilaterally [109].

The review literature in this field would also benefit substantially from more deliberate visual synthesis. A timeline of key methodological advances—marking the introduction of transfer learning to AM defect detection (approximately 2018), the first LSTM-based temporal monitoring demonstrations (2019), the arrival of physics-informed neural networks in AM contexts (2020–2021), and the emergence of federated learning frameworks (2022–2023)—would give readers, particularly those from manufacturing engineering backgrounds without deep ML familiarity, a coherent narrative of how the field has developed and where the current frontier sits. Equally valuable would be a comparative table of state-of-the-art model performance across defect types and AM processes, standardising the reporting of accuracy, F1-score, inference latency, dataset size, and whether cross-machine validation was performed. Such a table would immediately reveal the degree to which reported performance figures are dataset-specific rather than generalisable, making visible a structural pattern that is currently obscured by the diversity of experimental setups and reporting conventions across individual papers. Future review contributions in this space are strongly encouraged to invest in these synthesis tools, not merely as presentational aids but as analytical instruments that make the state of the field legible to the full audience—researchers, engineers, regulators, and technology managers—who need to act on it.

9.2. Physics-Informed and Hybrid Machine Learning

Integrating physical knowledge into neural network architectures—through PINNs, physics-constrained loss functions, thermodynamically consistent feature representations, and hybrid models coupling data-driven components with analytical process models [110,111]—represents the most scientifically rich frontier in ML-AM research. The promise of this approach extends beyond improved accuracy to improved reliability: a model architecturally constrained against physically inconsistent predictions is one whose failure modes are intrinsically better bounded and whose extrapolation behaviour beyond the training distribution is more defensible. Fully realising this potential requires deep, sustained collaboration among ML researchers, computational metallurgists, and AM process engineers—precisely the kind of genuinely interdisciplinary partnership that academic incentive structures are poorly designed to encourage, but that the complexity of this challenge demands [112,113]. Sajadi and colleagues [114] demonstrated one practical route forward, introducing the first physics-informed online learning framework for temperature prediction in metal AM; by continuously updating model weights as new process data arrive, their approach remained accurate for conditions the model had never seen during initial training—something static offline models routinely fail to do at scale.

9.3. Federated Learning: Breaking Down the Data Silos

Federated learning—a framework for collaborative model training across multiple institutions without exchanging raw data—offers an elegant technical solution to the data-sharing barriers currently fragmenting the AM-ML field. McMahan et al. established the foundational federated learning framework for communication-efficient training across decentralised data, demonstrating that global models trained without raw data exchange consistently outperform locally trained counterparts—a principle directly applicable to cross-machine AM defect detection. Each participating organisation trains a local model update on its proprietary build data and shares only the model weight gradients—not the underlying data—with a central aggregation server. In return, it receives a global model that has benefited from the collective experience of all participants. The confidentiality protection this architecture provides makes data participation acceptable to industrial organisations that would never agree to share raw build records. The medical imaging community has pioneered federated learning under closely analogous data sensitivity constraints; the metal AM field is well-positioned to adapt and build upon this experience while accounting for the specific challenges posed by the highly heterogeneous, non-identically distributed datasets generated by different AM machines and materials [115].

However, identifying federated learning as a technical priority and actually deploying it in industrial AM contexts are separated by a substantial gap that the roadmap must acknowledge honestly. Three categories of practical barrier stand between the current state and a functioning federated AM-ML ecosystem. The first is regulatory and legal. Build data from aerospace and defence AM suppliers is frequently subject to ITAR restrictions, export control classifications, and contractual non-disclosure obligations that may constrain even gradient sharing—since sufficiently detailed model weight updates can in principle be inverted to partially reconstruct training data. Navigating these constraints requires legal clarity that does not yet exist in a federated ML context, and engaging regulatory counsel proactively is a prerequisite for any industry-spanning initiative. The second barrier is economic and incentive-related. For a company that has invested years of process development in building a high-performing proprietary dataset, the federated model must offer a return that clearly exceeds what that company could achieve by training on its own data—otherwise participation is irrational from a competitive standpoint. Current evidence suggests that the federated advantage grows substantially with participant heterogeneity and diminishes when participants share similar machines and materials, which means that the organisations with the least to learn from each other are paradoxically the most likely to collaborate. Designing participation incentives—tiered access structures, weighted benefit sharing, or commercially negotiated data credit mechanisms—is an economic design problem that the technical community has not adequately addressed. The third barrier is technical and operational: achieving meaningful federation requires standardised data annotation protocols across participating organisations, which in turn demands agreement on defect taxonomies, labelling conventions, sensor calibration standards, and quality metric definitions that do not currently exist in a harmonised form across the industry.

Actionable paths forward exist for each of these barriers. Industry-academia consortia—structured along the lines of existing collaborative frameworks such as the Manufacturing Technology Centre in the UK, America Makes in the United States, or the Clean Sky 2 consortium in Europe—provide a governance model for pooling build data under negotiated intellectual property protections, with academic institutions acting as neutral data custodians operating under institutional data governance agreements. Standardised data annotation protocols can be developed through ASTM Committee F42 and ISO/TC 261 working groups, which already possess the convening authority and cross-sector membership required to establish consensus definitions of defect types, acceptable labelling methodologies, and minimum dataset documentation requirements. Addressing the participation incentive problem may ultimately require industry-side leadership: OEM primes and Tier 1 suppliers that have an economic interest in AM supply chain quality—and the purchasing leverage to drive supplier behaviour—are the most plausible catalysts for federating the supplier networks that generate the data the community needs. Naming these barriers explicitly in the roadmap, rather than treating federated learning as a straightforward technical proposition, is a necessary step toward mobilising the cross-sector engagement that will actually bring it to fruition.

9.4. Edge AI and Embedded Deployment

Closing the feedback loop between *in-situ* monitoring and real-time process control ultimately requires ML inference to execute at the edge—on or immediately adjacent to the machine controller—with latencies measured in milliseconds. The hardware ecosystem supporting edge ML inference has matured considerably: dedicated neural processing units embedded in the NVIDIA Jetson AGX Orin, the Hailo-8 accelerator, and Intel's Movidius Myriad X now deliver GPU-class inference performance within the power envelopes and physical form factors compatible with machine controller integration. The NVIDIA Jetson AGX Orin platform has been specifically validated for this class of embedded ML inference workload, delivering GPU-class performance within the power budgets and form factors compatible with commercial LPBF machine controller integration, enabling sub-50 ms inference latency. The accompanying software toolchains—TensorRT, OpenVINO, and TFLite with hardware delegates—automate the quantisation, pruning, and hardware-specific compilation steps needed to translate research-grade model implementations into optimised, edge-deployable forms. Research groups in this field should incorporate hardware-in-the-loop evaluation on representative edge platforms into their standard validation protocols from the outset, not as an afterthought [116].

9.5. Regulatory Engagement and Certification Pathways

Technical capability without regulatory acceptance translates into potential without impact. The AM-ML research community urgently needs to engage proactively with the regulatory bodies—the FAA, FDA, EASA, and their equivalents in defence procurement—whose decisions will determine whether ML-based monitoring evidence is admissible in part certification dossiers. This engagement should begin within the technical standards community: ASTM Committee F42 on Additive Manufacturing, ISO/TC 261, and their joint working groups are the natural venues for developing the performance specifications and validation requirements that ML monitoring systems must satisfy for regulatory acceptance. The specific technical outputs needed are clear: standardised uncertainty quantification methodologies analogous to the probability-of-detection (POD) curves that underpin conventional NDT qualification; protocols for evaluating ML model robustness to domain shift; and criteria defining minimum training dataset requirements and validation methodologies for ML models intended for regulated safety-critical manufacturing [117].

10. Outlook

Looking ahead, the convergence of several parallel technical developments is likely to reshape the landscape of ML-driven quality assurance for metal AM in ways that go well beyond incremental improvement on existing approaches. The following trends represent, in the authors' assessment, the most consequential directions that the field will need to navigate over the coming five to ten years.

Foundation models and transfer at scale. The emergence of large pre-trained foundation models in computer vision and multimodal learning opens a genuinely new possibility for AM defect detection: a single, massively pre-trained model that can be fine-tuned to a specific AM process, material, and sensor configuration with a fraction of the labelled data that current approaches require. Models in the GPT and CLIP lineage have already demonstrated remarkable cross-domain generalisation in natural image tasks; their adaptation to the structured, physics-governed imagery of AM monitoring is a near-term prospect rather than a distant aspiration. The critical research question is not whether such adaptation is possible but how much domain-specific data is needed to achieve industrially reliable performance, and whether the internal representations these models develop are physically meaningful or statistically coincidental—a distinction with profound implications for regulatory acceptance.

Physics-ML co-design as a standard rather than a speciality. Physics-informed neural networks and hybrid data-physics models are currently treated as a specialised subfield pursued by groups with both ML and computational materials expertise. Over the coming decade, the expectation is that physical constraint integration will become a standard design requirement rather than an optional enhancement—driven by regulatory pressure, by the interpretability demands of certification bodies, and by the demonstrated performance advantages of models that cannot make physically impossible predictions. This shift will require changes in how AM-ML researchers are trained: the next generation of practitioners will need fluency in both machine learning methodology and the thermomechanical physics of rapid solidification, a combination that current educational programmes rarely provide.

Closed-loop autonomous manufacturing systems. The trajectory of reinforcement learning and closed-loop adaptive control points toward AM machines that do not merely detect and report defects but actively prevent them—machines that learn from every build, accumulate process knowledge across materials and geometries, and autonomously optimise their own parameters within defined safety envelopes. Akmal and colleagues [118] offered early experimental evidence that this is not purely speculative: their AI-driven system for Ti-6Al-4V in laser powder bed fusion demonstrated that the process could self-heal defective regions—porosity introduced by deliberate parameter perturbations—within seven to eight subsequent layers under standard energy density conditions, a result that closes the loop between detection and correction in a physically measurable way. The realisation of this vision depends critically on advances in sim-to-real transfer, on the development of high-fidelity process simulators that can serve as training environments for RL agents, and on the establishment of safety frameworks that define the bounds within which autonomous parameter adjustment is permissible without human oversight. The latter is as much a regulatory and organisational challenge as a technical one, and its resolution will require dialogue between the ML research community, machine manufacturers, and the aerospace and medical device certification authorities whose sign-off will ultimately determine whether autonomous AM is commercially viable.

Democratisation through edge AI and cloud-connected monitoring platforms. As edge inference hardware matures and cloud-connected monitoring platforms become commercially viable, the capability gap between well-resourced research institutions and small-to-medium manufacturers will narrow. The implications are significant: ML-based quality assurance may within a decade become accessible to the mid-market AM service bureaux and job shops that currently lack the data science expertise and capital to implement bespoke monitoring systems. This democratisation will expand the pool of contributors to federated learning networks, diversify the machine and material combinations represented in community datasets, and accelerate the accumulation of real-world evidence that regulatory frameworks require. The challenge will be ensuring that accessibility does not come at the cost of rigour—that off-the-shelf monitoring solutions are validated to the same evidential standards as custom research implementations.

Human-machine collaboration and operator-centred design. A thread running through the technical optimism of this field that deserves more explicit attention is the role of the human operator. The most capable ML monitoring system deployed in isolation from the expertise of experienced AM technicians will underperform relative to one designed as a collaborative tool—surfacing model uncertainty to the operator, making its reasoning legible through effective visualisation, and deferring to human judgement in the ambiguous cases that no model handles with complete confidence. Human factors research in AM quality assurance is currently sparse relative to algorithm development; this imbalance should be corrected as the field moves from demonstration to deployment. Systems that operators trust, understand, and can effectively override will be adopted far more rapidly than systems of equivalent or superior technical performance that treat the human as a passive output recipient.

Taken together, these trends suggest that the next decade will see ML-driven AM quality assurance transition from a research frontier to an industrial standard—not uniformly, and not without setbacks, but with a trajectory that the community can meaningfully shape through deliberate investment in the shared infrastructure, the regulatory engagement, and the cross-disciplinary collaboration that the challenge demands. The technical foundations are stronger than they have ever been. What remains is to build the ecosystem around them.

11. Conclusions

This review has examined the state of machine learning for process optimisation and defect detection in metal additive manufacturing with both genuine appreciation for the progress made and rigorous scrutiny of its limitations. The enthusiasm is well-founded: the past five years have yielded substantive advances across the field. CNNs combined with transfer learning have demonstrated that image-based defect detection approaching or exceeding human-level accuracy is achievable from *in-situ* optical data. LSTM and transformer architectures have established that the temporal dynamics of the build process are learnable in ways that enable predictive rather than merely reactive quality assessment. Closed-loop reinforcement learning has advanced from simulation-only demonstrations to physical hardware trials with measurable quality improvements. Physics-informed neural networks have made real-time digital twin updates computationally feasible for the first time. And the commercial ecosystem of *in-situ* monitoring hardware and software has matured to the point where industrial practitioners have meaningful choices in how they instrument their machines.

The critical scrutiny is equally warranted. Performance metrics reported on small, homogeneous, single-machine datasets systematically overstate real-world deployability, in some cases substantially. Domain shift across machines, materials, and build conditions is a pervasive and persistently underacknowledged failure mode that invalidates the generalisation claims of many published models. The data shortage is a structural constraint rather than an incidental one, and it cannot be resolved through individual laboratory efforts alone. The computational demands of real-time embedded deployment are more severe than most papers acknowledge, and more frequently deferred than addressed. The regulatory pathway from compelling research demonstration to certified industrial quality assurance system remains long, demanding, and inadequately mapped by the community working to navigate it.

The organisational and product development challenges associated with AM adoption compound the technical ones: Seifi et al. [119] documented the standardisation progress needed to support qualification and certification of metal AM parts, identifying the specific regulatory gaps that must be bridged before ML-based monitoring evidence can be incorporated into part qualification dossiers—a prerequisite for the systematic quality assurance tools reviewed here. Emerging applications of ML in AM are also extending beyond process monitoring into downstream quality attributes: Mani et al. [120] identified the measurement science needs for real-time control of AM powder bed fusion processes, establishing the sensor and data requirements that ML-based quality assurance systems must satisfy to support industrially deployable closed-loop control across dimensional and defect quality attributes. The researchers, engineers, and institutions working on these problems are not deficient in intelligence or creative ambition. What they frequently lack are the shared data infrastructure, the cross-disciplinary collaboration structures, and the proactive regulatory engagement that would allow the best ideas generated in individual laboratories to accumulate into durable, trustworthy, and certifiable systems. Building those shared foundations—the community benchmark datasets, the federated data networks, the technical standards frameworks, the sustained engineer-regulator dialogues—is less glamorous than developing a novel neural architecture or reporting a record accuracy figure. It is, however, the work upon which the long-term industrial value of this entire research programme fundamentally depends. This review is offered in the belief that naming these needs clearly, honestly, and specifically will contribute to accelerating the coordinated community action required to meet them.

Author Contributions

Conceptualization: I.E. and D.J.; Methodology and formal analysis: I.E. and A.I.; Investigation and data curation: D.J. and S.C.A.; Validation: A.U.M.; Writing—original draft: I.E.; Writing—review and editing: all authors; Supervision: I.E.; Funding acquisition: I.E. and S.C.A. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data Availability Statement

No new data were created or analysed in this review study. Data sharing is not applicable to this article.

Acknowledgments

The authors acknowledge the Department of Mechanical Engineering, Nnamdi Azikiwe University, Awka, Nigeria, for providing the enabling research environment. This manuscript was prepared in full accordance with the author guidelines of the Journal of Mechanical Engineering and Manufacturing.

Conflicts of Interest

The authors declare no conflict of interest.

Use of AI and AI-Assisted Technologies

No AI tools were utilized for this paper.

References

1. Attaran, M. The rise of 3D printing: The advantages of additive manufacturing over traditional manufacturing. *Bus. Horiz.* **2017**, *60*, 677–688.
2. SAE International. *AS9100 Rev. D: Quality Management Systems—Requirements for Aviation, Space, and Defense Organizations*; SAE International: Warrendale, PA, USA, 2016.
3. International Organization for Standardization. *ISO 13485:2016—Medical Devices: Quality Management Systems*; ISO: Geneva, Switzerland, 2016.
4. Tapia, G.; Elwany, A. A review on process monitoring and control in metal-based additive manufacturing. *J. Manuf. Sci. Eng.* **2014**, *136*, 060801.
5. Taherkhani, K.; Ero, O.; Liravi, F.; et al. On the application of *in-situ* monitoring systems and machine learning algorithms for developing quality assurance platforms in laser powder bed fusion: A review. *J. Manuf. Process.* **2023**, ***99***, 848–897.
6. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* **2021**, *372*, n71.
7. Liberati, A.; Altman, D.G.; Tetzlaff, J.; et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions. *PLoS Med.* **2009**, *6*, e1000100.
8. Moher, D.; Liberati, A.; Tetzlaff, J.; et al. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Ann. Intern. Med.* **2009**, *151*, 264–269.
9. Kitchenham, B.; Charters, S. *Guidelines for Performing Systematic Literature Reviews in Software Engineering*; Technical Report EBSE-2007-01; Keele University and Durham University: Staffordshire, UK, 2007.
10. DebRoy, T.; Wei, H.L.; Zuback, J.S.; et al. Additive manufacturing of metallic components—Process, structure and properties. *Prog. Mater. Sci.* **2018**, *92*, 44–224.
11. King, W.E.; Anderson, A.T.; Ferencz, R.M.; et al. Laser powder bed fusion additive manufacturing of metals: Physics, computational, and materials challenges. *Appl. Phys. Rev.* **2015**, *2*, 041304.
12. Khairallah, S.A.; Anderson, A.T.; Rubenchik, A.; et al. Laser powder-bed fusion additive manufacturing: Physics of complex melt flow and formation mechanisms of pores, spatter, and denudation zones. *Acta Mater.* **2016**, *108*, 36–45.
13. Gao, W.; Zhang, Y.; Ramanujan, D.; et al. Additive manufacturing: Technology, applications and research needs. *Front. Mech. Eng.* **2013**, *10*, 215–243.
14. Thompson, S.M.; Bian, L.; Shamsaei, N.; et al. An overview of Direct Laser Deposition for additive manufacturing; Part I: Transport phenomena, modeling and diagnostics. *Addit. Manuf.* **2015**, *8*, 36–62.
15. Juechter, V.; Scharowsky, T.; Singer, R.F.; et al. Processing window and evaporation phenomena for Ti–6Al–4V produced by selective electron beam melting. *Acta Mater.* **2014**, *76*, 252–258.
16. Sames, W.J.; List, F.A.; Pannala, S.; et al. The metallurgy and processing science of metal additive manufacturing. *Int. Mater. Rev.* **2016**, *61*, 315–360.
17. Leung, C.L.A.; Marussi, S.; Atwood, R.C.; et al. *In situ* X-ray imaging of defect and molten pool dynamics in laser additive manufacturing. *Nat. Commun.* **2018**, *9*, 1355.
18. Cunningham, R.; Zhao, C.; Parab, N.; et al. Keyhole threshold and morphology in laser melting revealed by ultrahigh-speed x-ray imaging. *Science* **2019**, *363*, 849–852.
19. Huang, Y.; Fleming, T.G.; Clark, S.J.; et al. Keyhole fluctuation and pore formation mechanisms during laser powder bed fusion additive manufacturing. *Nat. Commun.* **2022**, *13*, 1170.
20. Carter, L.N.; Martin, C.; Withers, P.J.; et al. The influence of the laser scan strategy on grain structure and cracking behaviour in SLM powder-bed fabricated nickel superalloy. *J. Alloys Compd.* **2014**, *615*, 338–347.

21. Zhao, C.; Fezzaa, K.; Cunningham, R.W.; et al. Real-time monitoring of laser powder bed fusion process using high-speed X-ray imaging and diffraction. *Sci. Rep.* **2017**, *7*, 3602.
22. Mercelis, P.; Kruth, J.P. Residual stresses in selective laser sintering and selective laser melting. *Rapid Prototyp. J.* **2006**, *12*, 254–265.
23. Mukherjee, T.; Zuback, J.S.; De, A.; et al. Printability of alloys for additive manufacturing. *Sci. Rep.* **2016**, *6*, 19717.
24. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444.
25. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. *Commun. ACM* **2017**, *60*, 84–90.
26. Scime, L.; Beuth, J. Anomaly detection and classification in a laser powder bed additive manufacturing process using a trained computer vision algorithm. *Addit. Manuf.* **2018**, *19*, 114–126.
27. Caggiano, A.; Zhang, J.; Alfieri, V.; et al. Machine learning-based image processing for on-line defect recognition in additive manufacturing. *CIRP Ann.* **2019**, *68*, 451–454.
28. Gobert, C.; Reutzel, E.W.; Petrich, J.; et al. Application of supervised machine learning for defect detection during metallic powder bed fusion additive manufacturing using high resolution imaging. *Addit. Manuf.* **2018**, *21*, 517–528.
29. Kwon, O.; Kim, H.G.; Ham, M.J.; et al. A deep neural network for classification of melt-pool images in metal additive manufacturing. *J. Intell. Manuf.* **2020**, *31*, 375–386.
30. Tan, M.; Le, Q. EfficientNet: Rethinking model scaling for convolutional neural networks. In Proceedings of the 36th International Conference on Machine Learning, Long Beach, CA, USA, 9–15 June 2019.
31. Weiss, K.; Khoshgoftaar, T.M.; Wang, D. A survey of transfer learning. *J. Big Data* **2016**, *3*, 9.
32. Pandiyan, V.; Drissi-Daoudi, R.; Shevchik, S.; et al. Deep transfer learning of additive manufacturing mechanisms across materials in metal-based laser powder bed fusion process. *J. Mater. Process. Technol.* **2022**, *303*, 117531.
33. Jafari-Marandi, R.; Khanzadeh, M.; Tian, W.; et al. From *in-situ* monitoring toward high-throughput process control: Cost-driven decision-making framework for laser-based additive manufacturing. *J. Manuf. Syst.* **2019**, *51*, 29–41.
34. Westphal, E.; Seitz, H. A machine learning method for defect detection and visualization in selective laser sintering based on convolutional neural networks. *Addit. Manuf.* **2021**, *41*, 101965.
35. Zhang, Y.; Soon, H.G.; Ye, D.; et al. Extraction and evaluation of melt pool, plume and spatter information for powder-bed fusion AM process monitoring. *Mater. Des.* **2018**, *156*, 458–469.
36. Scime, L.; Siddel, D.; Baird, S.; et al. Layer-wise anomaly detection and classification for powder bed additive manufacturing processes: A machine-agnostic algorithm for real-time pixel-wise semantic segmentation. *Addit. Manuf.* **2020**, *36*, 101453.
37. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780.
38. Ye, D.; Hong, G.S.; Zhang, Y.; et al. Defect detection in selective laser melting technology by acoustic signals with deep belief networks. *Int. J. Adv. Manuf. Technol.* **2018**, *96*, 2791–2801.
39. Khanzadeh, M.; Chowdhury, S.; Marufuzzaman, M.; et al. Porosity prediction: Supervised-learning of thermal history for direct laser deposition. *J. Manuf. Syst.* **2018**, *47*, 69–82.
40. Williams, R.J.; Davies, C.M.; Hooper, P.A. A pragmatic part scale model for residual stress and distortion prediction in powder bed fusion. *Addit. Manuf.* **2018**, *22*, 416–425.
41. Dosovitskiy, A.; Beyer, L.; Kolesnikov, A.; et al. An image is worth 16×16 words: Transformers for image recognition at scale. In Proceedings of the ICLR 2021, Vienna, Austria, 4 May 2021.
42. Sutton, R.S.; Barto, A.G. *Reinforcement Learning: An Introduction*, 2nd ed.; MIT Press: Cambridge, MA, USA, 2018.
43. Mnih, V.; Kavukcuoglu, K.; Silver, D.; et al. Human-level control through deep reinforcement learning. *Nature* **2015**, *518*, 529–533.
44. Donadello, I.; Motta, M.; Demir, A.G.; et al. Monitoring of laser metal deposition height by means of coaxial laser triangulation. *Opt. Lasers Eng.* **2019**, *112*, 136–144.
45. Schulman, J.; Wolski, F.; Dhariwal, P.; et al. Proximal policy optimization algorithms. *arXiv* **2017**, arXiv:1707.06347.
46. Haarnoja, T.; Zhou, A.; Abbeel, P.; et al. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. In Proceedings of the 35th International Conference on Machine Learning, Stockholm, Sweden, 10–15 July 2018.
47. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; et al. Generative adversarial nets. In Proceedings of the Annual Conference on Neural Information Processing Systems 2014, Montreal, QC, Canada, 8–13 December 2014.
48. Ho, J.; Jain, A.; Abbeel, P. Denoising diffusion probabilistic models. In Proceedings of the Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, Virtual, 6–12 December 2020.
49. Kingma, D.P.; Welling, M. Auto-encoding variational Bayes. In Proceedings of the ICLR 2014, Banff, AB, Canada, 14–16 April 2014.
50. Paul, R.; Anand, S.; Gerner, F. Effect of thermal deformation on part errors in metal powder based additive manufacturing processes. *J. Manuf. Sci. Eng.* **2014**, *136*, 031009.

51. Ero, O.; Taherkhani, K.; Toyserkani, E. Optical tomography and machine learning for *in-situ* defect detection in laser powder bed fusion: A self-organizing map and U-Net based approach. *Addit. Manuf.* **2023**, ***78***, 103894.
52. Everton, S.K.; Hirsch, M.; Stravroulakis, P.; et al. Review of *in-situ* process monitoring and *in-situ* metrology for metal additive manufacturing. *Mater. Des.* **2016**, *95*, 431–445.
53. Grasso, M.; Colosimo, B.M. Process defects and *in situ* monitoring methods in metal powder bed fusion: A review. *Meas. Sci. Technol.* **2017**, *28*, 044005.
54. Snow, Z.; Nassar, A.R.; Reutzel, E.W. Invited review article: Review of the formation and impact of flaws in powder bed fusion additive manufacturing. *Addit. Manuf.* **2020**, ***36***, 101457.
55. Grasso, M.; Laguzza, V.; Semeraro, Q.; et al. In-process monitoring of selective laser melting: Spatial detection of defects via image processing. *J. Manuf. Sci. Eng.* **2017**, *139*, 051001.
56. Clijsters, S.; Craeghs, T.; Buls, S.; et al. *In situ* quality control of the selective laser melting process using a high-speed, real-time melt pool monitoring system. *Int. J. Adv. Manuf. Technol.* **2014**, *75*, 1089–1101.
57. Imani, F.; Gaikwad, A.; Montazeri, M.; et al. Process mapping and in-process monitoring of porosity in laser powder bed fusion using layerwise optical imaging. *J. Manuf. Sci. Eng.* **2018**, *140*, 101009.
58. Spears, T.G.; Gold, S.A. In-process sensing in selective laser melting (SLM) additive manufacturing. *Integr. Mater. Manuf. Innov.* **2016**, *5*, 16–40.
59. Hooper, P.A. Melt pool temperature and cooling rates in laser powder bed fusion. *Addit. Manuf.* **2018**, *22*, 548–559.
60. Wolff, S.J.; Webster, S.; Parab, N.D.; et al. *In-situ* high-speed X-ray imaging of piezo-driven directed energy deposition additive manufacturing. *Sci. Rep.* **2019**, *9*, 962.
61. Oster, S.; Breese, P.P.; Ulbricht, A.; et al. A deep learning framework for defect prediction based on thermographic *in-situ* monitoring in laser powder bed fusion. *J. Intell. Manuf.* **2024**, ***35***, 1687–1706.
62. Krauss, H.; Zeugner, T.; Zaeh, M.F. Layerwise monitoring of the selective laser melting process by thermography. *Phys. Procedia* **2014**, *56*, 64–71.
63. Shevchik, S.A.; Kenel, C.; Leinenbach, C.; et al. Acoustic emission for *in situ* quality monitoring in additive manufacturing using spectral convolutional neural networks. *Addit. Manuf.* **2018**, *21*, 598–604.
64. Wasmer, K.; Kenel, C.; Leinenbach, C.; et al. *In situ* quality monitoring in AM using acoustic emission: A reinforcement learning approach. *J. Mater. Eng. Perform.* **2019**, *28*, 666–672.
65. Repossini, G.; Laguzza, V.; Grasso, M.; et al. On the use of spatter signature for *in-situ* monitoring of laser powder bed fusion. *Addit. Manuf.* **2017**, *16*, 35–48.
66. Mahmoudi, M.; Elwany, A.; Yadollahi, A.; et al. Mechanical properties and microstructural characterization of selective laser melted 17-4 PH stainless steel. *Rapid Prototyp. J.* **2017**, *23*, 280–294.
67. Stavridis, J.; Papacharalampopoulos, A.; Stavropoulos, P. Recent advances in sensor fusion monitoring and control strategies in laser powder bed fusion: A review. *Machines* **2025**, ***13***, 820.
68. Petrich, J.; Snow, Z.; Corbin, D.; et al. Multi-modal sensor fusion with machine learning for data-driven process monitoring for additive manufacturing. *Addit. Manuf.* **2021**, *48*, 102364.
69. Shen, N.; Chou, K. Simulations of thermo-mechanical characteristics in electron beam additive manufacturing. In Proceedings of the ASME 2012 International Mechanical Engineering Congress and Exposition, Houston, TX, USA, 9–15 November 2012.
70. Baumgartl, H.; Tomas, J.; Buettner, R.; et al. A deep learning-based model for defect detection in laser-powder bed fusion using *in-situ* thermographic monitoring. *Prog. Addit. Manuf.* **2020**, *5*, 277–285.
71. Wang, P.; Yang, Y.; Moghaddam, N.S. Process modeling in laser powder bed fusion towards defect detection and quality control via machine learning: The state-of-the-art and research challenges. *J. Manuf. Process.* **2022**, *73*, 961–984.
72. Yeung, H.; Lane, B.; Fox, J. Part geometry and conduction-based laser power control for powder bed fusion additive manufacturing. *Addit. Manuf.* **2019**, *30*, 100844.
73. Chen, L.; Bi, G.; Yao, X.; et al. *In-situ* process monitoring and adaptive quality enhancement in laser additive manufacturing: A critical review. *J. Manuf. Syst.* **2024**, ***74***, 527–574.
74. Mondal, S.; Gwynn, D.; Ray, A.; et al. Investigation of melt pool geometry control in additive manufacturing using hybrid modeling. *Metals* **2020**, *10*, 683.
75. Mazzucato, F.; Tusacciu, S.; Lai, M.; et al. Monitoring approach to evaluate the performances of a new deposition nozzle solution for DED additive manufacturing. *Technologies* **2017**, *5*, 29.
76. Mozaffar, M.; Paul, A.; Al-Bahrani, R.; et al. Data-driven prediction of the high-dimensional thermal history in directed energy deposition processes. *Manuf. Lett.* **2018**, *18*, 35–39.
77. Renken, V.; von Freyberg, A.; Schumann, K.; et al. In-process closed-loop control for stabilising the melt pool temperature in selective laser melting. *Prog. Addit. Manuf.* **2019**, *4*, 411–421.
78. Grieves, M.; Vickers, J. Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems. In *Transdisciplinary Perspectives on Complex Systems*; Springer: Cham, Switzerland, 2017; pp. 85–113.

79. Lu, Y.; Liu, C.; Kevin, I.; et al. Digital twin-driven smart manufacturing: Connotation, reference model, applications and research issues. *Robot. CIM* **2020**, *61*, 101837.
80. Tudorache, L.; Babur, Ö.; Lucas, S.S.; et al. Current approaches to digital twins in additive manufacturing: A systematic literature review. *Prog. Addit. Manuf.* **2025**, ***10***, 10819–10853.
81. Nath, P.; Mahadevan, S. Probabilistic digital twin for additive manufacturing process design and control. *J. Mech. Des.* **2022**, ***144***, 091703.
82. Raissi, M.; Perdikaris, P.; Karniadakis, G.E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* **2019**, *378*, 686–707.
83. Karniadakis, G.E.; Kevrekidis, I.G.; Lu, L.; et al. Physics-informed machine learning. *Nat. Rev. Phys.* **2021**, *3*, 422–440.
84. Liu, R.; Liu, S.; Zhang, X. A physics-informed machine learning model for porosity analysis in laser powder bed fusion additive manufacturing. *Int. J. Adv. Manuf. Technol.* **2021**, *113*, 1943–1958.
85. Jiang, F.; Xia, M.; Hu, Y. Physics-informed machine learning for accurate prediction of temperature and melt pool dimension in metal additive manufacturing. *3D Print. Addit. Manuf.* **2024**, ***11***, e1679–e1689.
86. Zhu, Q.; Liu, Z.; Yan, J. Machine learning for metal additive manufacturing: Predicting temperature and melt pool fluid dynamics using physics-informed neural networks. *Comput. Mech.* **2021**, *67*, 619–635.
87. ISO/ASTM International. *ISO/ASTM 52941:2025—Additive Manufacturing—System Performance and Reliability—Acceptance Tests for Laser Metal Powder Bed Fusion Machines for Metallic Materials for Aerospace Applications*; ISO: Geneva, Switzerland, 2025.
88. ASTM International. F3187-16: Standard Guide for Directed Energy Deposition of Metals; ASTM International: West Conshohocken, PA, USA, 2016.
89. ISO/ASTM International. *ISO/ASTM 52920:2023—Additive Manufacturing—Qualification Principles*; ISO: Geneva, Switzerland, 2023.
90. Adadi, A.; Berrada, M. Peeking inside the black-box: A survey on explainable AI. *IEEE Access* **2018**, *6*, 52138–52160.
91. Oliveira, J.P.; LaLonde, A.D.; Ma, J. Processing parameters in laser powder bed fusion metal additive manufacturing. *Mater. Des.* **2020**, *193*, 108762.
92. Gu, D.; Shi, X.; Poprawe, R.; et al. Material-structure-performance integrated laser-metal additive manufacturing. *Science* **2021**, *372*, eabg1487.
93. Wang, C.; Tan, X.P.; Tor, S.B.; et al. Machine learning in additive manufacturing: State-of-the-art and perspectives. *Addit. Manuf.* **2020**, *36*, 101538.
94. Qi, X.; Chen, G.; Li, Y.; et al. Applying neural-network-based machine learning to additive manufacturing: Current applications, challenges, and future perspectives. *Engineering* **2019**, *5*, 721–729.
95. Gan, Z.; Jones, K.K.; Lu, Y.; et al. Benchmark study of melted track geometries in laser powder bed fusion of Inconel 625. *Integr. Mater. Manuf. Innov.* **2021**, *10*, 177–195.
96. Tapia, G.; Khairallah, S.; Matthews, M.; et al. Gaussian process-based surrogate modeling framework for process planning in laser powder-bed fusion additive manufacturing of 316L stainless steel. *Int. J. Adv. Manuf. Technol.* **2018**, *94*, 3591–3603.
97. Brochu, E.; Cora, V.M.; de Freitas, N. A tutorial on Bayesian optimization of expensive cost functions. *arXiv* **2010**, arXiv:1012.2599.
98. Snoek, J.; Larochelle, H.; Adams, R.P. Practical Bayesian optimization of machine learning algorithms. In Proceedings of the 26th Annual Conference on Neural Information Processing Systems 2012, Lake Tahoe, NV, USA, 3–6 December 2012.
99. Shahriari, B.; Swersky, K.; Wang, Z.; et al. Taking the human out of the loop: A review of Bayesian optimization. *Proc. IEEE* **2016**, *104*, 148–175.
100. Hertlein, N.; Deshpande, S.; Venugopal, V.; et al. Prediction of selective laser melting part quality using hybrid Bayesian network. *Addit. Manuf.* **2020**, *32*, 101089.
101. Okaro, I.A.; Jayasinghe, S.; Sutcliffe, C.; et al. Automatic fault detection for laser powder-bed fusion using semi-supervised machine learning. *Addit. Manuf.* **2019**, *27*, 42–53.
102. Meng, L.; McWilliams, B.; Jarosinski, W.; et al. Machine learning in additive manufacturing: A review. *JOM* **2020**, *72*, 2363–2377.
103. Johnson, N.S.; Vulimiri, P.S.; To, A.C.; et al. Invited review: Machine learning for materials developments in metals additive manufacturing. *Addit. Manuf.* **2020**, *36*, 101641.
104. Gibson, I.; Rosen, D.; Stucker, B.; et al. *Additive Manufacturing Technologies*, 3rd ed.; Springer: New York, NY, USA, 2021.
105. Wilson, A.G.; Hu, Z.; Salakhutdinov, R.; et al. Deep kernel learning. In Proceedings of the 19th International Conference on Artificial Intelligence and Statistics, Cadiz, Spain, 9–11 May 2016.
106. Han, S.; Pool, J.; Tran, J.; et al. Learning both weights and connections for efficient neural network. In Proceedings of the Annual Conference on Neural Information Processing Systems 2015, Montreal, QC, Canada, 7–12 December 2015.
107. Hinton, G.; Vinyals, O.; Dean, J. Distilling the knowledge in a neural network. *arXiv* **2015**, arXiv:1503.02531.

108. Guo, S.; Agarwal, M.; Cooper, C.; et al. Machine learning for metal additive manufacturing: Towards a physics-informed data-driven paradigm. *J. Manuf. Syst.* **2022**, ***62***, 145–163.
109. Deng, J.; Dong, W.; Socher, R.; et al. ImageNet: A large-scale hierarchical image database. In Proceedings of the IEEE CVPR 2009, Miami, FL, USA, 20–25 June 2009.
110. Psichogios, D.C.; Ungar, L.H. A hybrid neural network-first principles approach to process modeling. *AIChE J.* **1992**, *38*, 1499–1511.
111. Willard, J.; Jia, X.; Xu, S.; et al. Integrating scientific knowledge with machine learning for engineering and environmental systems. *ACM Comput. Surv.* **2022**, *55*, 1–37.
112. Bayat, M.; Dong, W.; Thorborg, J.; et al. A review of multi-scale and multi-physics simulations of metal additive manufacturing processes with focus on modeling strategies. *Addit. Manuf.* **2021**, *47*, 102278.
113. Loh, G.H.; Pei, E.; Harrison, D.; et al. An overview of functionally graded additive manufacturing. *Addit. Manuf.* **2018**, *23*, 34–44.
114. Sajadi, P.; Rahmani Dehaghani, M.; Tang, Y.; et al. Physics-informed online learning for temperature prediction in metal additive manufacturing. *Materials* **2024**, ***17***, 3306.
115. McMahan, H.B.; Moore, E.; Ramage, D.; et al. Communication-efficient learning of deep networks from decentralized data. In Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS) 2017, Fort Lauderdale, FL, USA, 20–22 April 2017.
116. NVIDIA Corporation. *NVIDIA Jetson AGX Orin Technical Brief*; NVIDIA: Santa Clara, CA, USA, 2022.
117. U.S. Food and Drug Administration. *Technical Considerations for Additive Manufactured Medical Devices*; FDA: Silver Spring, MD, USA, 2017.
118. Akmal, J.; Minet-Lallemand, K.; Kuva, J.; et al. AI-based defect detection and self-healing in metal additive manufacturing. *Virtual Phys. Prototyp.* **2025**, ***20***, e2500671.
119. Seifi, M.; Salem, A.; Beuth, J.; et al. Progress towards metal additive manufacturing standardization to support qualification and certification. *JOM* **2017**, *69*, 439–455.
120. Mani, M.; Lane, B.M.; Donmez, M.A.; et al. A review on measurement science needs for real-time control of additive manufacturing metal powder bed fusion processes. *Int. J. Prod. Res.* **2017**, *55*, 1400–1418.